Variational and Ergodic Methods for Stochastic Differential Equations Driven by Lévy Processes

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

im Fach Mathematik

eingereicht an der

Mathematisch Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von

Herrn Dipl. Math. Jan Martin Gairing

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Elmar Kulke

Gutachter:
1. Prof. Dr. Peter Imkeller
2. Prof. Dr. Michael Scheutzow (Technische Universität Berlin)
3. Prof. Dr. Martin Rasmussen (Imperial College London)

eingereicht am: 11.07.2017
Tag der mündlichen Prüfung: 12.01.2018

Nymphalis thamyris (Nymphalis iphis)

Alexander von Humboldt & Aimé Bonplant: *Recueil d'observation de zoologie et d'anatomie comparée, Deuxième Partie*, Paris 1821, Pl. XXXVI [Papillons].
(source: gallica.bnf.fr / Bibliothèque nationale de France)

Acknowledgements: The author thanks the DFG IRTG 1740: *Dynamical Phenomena in Complex Networks*, Berlin–São Paulo for financial support, the *Instituto de Matemática, Estatística e Computação Científica* (IMECC) at the University of Campinas (UNICAMP), in particular Prof. Dr. Paulo Ruffino, for the kind hospitality, and Prof. Dr. Peter Imkeller for his support and patience.

Summary

The present thesis investigates certain aspects of the interplay between the ergodic long time behavior and the smoothing property of dynamical systems generated by stochastic differential equations (*SDE*s) with jumps, in particular *SDE*s driven by Lévy processes and the Marcus' canonical equation. A variational approach to the Malliavin calculus generates an integration-by-parts formula that allows to transfer spatial variation to variation in the probability measure. The strong Feller property of the associated Markov semigroup and the existence of smooth transition densities are deduced from a non-standard ellipticity condition on a combination of the Gaussian and a jump covariance. Similar results on submanifolds are inferred from the ambient Euclidean space.

These results are then applied to random dynamical systems generated by linear stochastic differential equations. Ruelle's integrability condition translates into an integrability condition for the Lévy measure and ensures the validity of the multiplicative ergodic theorem (MET) of Oseledets. Hence the exponential growth rate is governed by the Lyapunov spectrum. Finally the top Lyapunov exponent is represented by a formula of Furstenberg–Khasminskii–type as an ergodic average of the infinitesimal growth rate over the unit sphere.

Zusammenfassung

Diese Dissertation untersucht Aspekte des Zusammenspiels von ergodischem Langzeitverhalten und der Glättungseigenschaft dynamischer Systeme, die von stochastischen Differentialgleichungen (*SDE*s) mit Sprüngen erzeugt sind. Im Speziellen werden *SDE*s getrieben von Lévy-Prozessen und der Marcusschen kanonischen Gleichung untersucht. Ein variationeller Ansatz für den Malliavin-Kalkül liefert eine partielle Integration, sodass eine Variation im Raum in eine Variation im Wahrscheinlichkeitsmaß überführt werden kann. Damit lässt sich die starke Feller-Eigenschaft und die Existenz glatter Dichten der zugehörigen Markov-Halbgruppe aus einer nichtstandard Elliptizitätsbedingung an eine Kombination aus Gaußscher und Sprung-Kovarianz ableiten. Resultate für Sprungdiffusionen auf Untermannigfaltigkeiten werden aus dem umgebenden Euklidischen Raum hergeleitet. Diese Resultate werden dann auf zufällige dynamische Systeme angewandt, die von linearen stochastischen Differentialgleichungen erzeugt sind. Ruelles Integrierbarkeitsbedingung entspricht einer Integrierbarkeitsbedingung an das Lévy-Maß und gewährleistet die Gültigkeit von Oseledets multiplikativem Ergodentheorem. Damit folgt die Existenz eines Lyapunov-Spektrums. Schließlich wird der top Lyapunov-Exponent über eine Formel der Art von Furstenberg–Khasminsikii als ein ergodisches Mittel der infinitesimalen Wachstumsrate über die Einheitssphäre dargestellt.

Meiner Familie gewidmet

Contents

Introduction

The present thesis investigates certain aspects of the interplay between the long time behavior and smoothing properties of dynamical systems generated by stochastic differential equations (*SDE*s) with jumps, in particular *SDE*s driven by Lévy processes. While in general processes are considered on the Euclidean space we also infer results for processes on compact submanifolds.
The work consists of two parts that are relatively independent although motivated through the fact that Part II relies on results obtained in Part I.
Part I investigates the regularity of the Markov semigroup generated by the *SDE* by means of the *stochastic calculus of variations* or *Malliavin calculus*.
Two main features constitute this regularity.
One important feature is the absolute continuity of the transition probability of the Markov process with respect to the Lebesgue measure, i.e. the existence of transition densities. This can be thought of as the differentiability of the Markov semigroup with respect to the terminal position.
The other in some way complementary feature is the smoothing property provided by the so called strong Feller property. This property can be verified if the transition probabilities are differentiable with respect to the initial condition.
Both properties are derived by means of an *integration-by-parts* formula which allows to transfer the variation in space to a variation of the probability measure under a certain ellipticity assumption. The ellipticity under consideration stems from a non-standard combination of the Gaussian and the jump covariance and hence does not require the ellipticity of the second order differential operator associated with the generator of the semigroup. The regularity of the semigroup is intimately related to the unique ergodicity of the corresponding Markov process in the sense that it provides the necessary topological irreducibility.
Part II focuses on exponential growth rates, called *Lyapunov exponents*, of linear dynamical systems that arise from linear versions of such *SDE*s. The main goal of this analysis is the expression of the top (i.e. the largest) Lyapunov exponent by an ergodic average of the infinitesimal growth rate over all directions represented by the unit sphere. The resulting formula will be called a formula of *Furstenberg–Khasminskii* type after the fundamental results of [Fur63] and [Kha67]. It is here that we can rely on the existence of a unique ergodic measure on the sphere due to results of Part I. The two parts are organized as follows.

Part I.

Chapter 1 starts with a short review on the canonical Wiener–Poisson probability space and its relation to Lévy processes. Then the concept of semimartingales with spatial parameter and *SDE*s based on it is presented. In particular Lévy driven *SDE*s (of multiplicative type) and the Marcus' canonical equation are reviewed.
Chapter 2 introduces the concept of L^p-derivatives. The concept is then applied to define the "Jacobian" of the stochastic flow generated by the *SDE* and to obtain flows of diffeomorphisms. In a next step the concept is applied to define the variation of the probability measure and to define directional derivatives of solutions to *SDE*s in the probability space.
Chapter 3 is devoted to the derivation of the integration-by-parts formula. The integration-by-parts formula enables us to define a dual operator to the L^p-derivative that is closable under a certain norm. The chapter closes with abstract criteria for the strong Feller property and the existence of smooth densities based on the integration-by-parts formula.
Chapter 4 investigates *Bismut–Elworthy–Li* formulae, i.e. the differential of the semigroup in the case where the Gaussian and the jump covariances are individually elliptic.
Chapter 5 considers the case where neither the Gaussian nor the jump covariance is elliptic on its own. Instead we require the sum of the two to be elliptic. With the help of the criteria developed in Chapter 3 we can prove the strong Feller property via gradient estimates and the existence of smooth densities.
Chapter 6 extends the previous results to jump diffusions on submanifolds of the Euclidean space. First a criterion for the trajectories to be confined to a submanifold is established. Then gradient estimates and the existence of densities are established under weaker conditions than the ellipticity in the ambient space.

Part II.

The second part analyses the exponential growth rate of linear Lévy driven versions of the *SDE* considered in Part I.
Chapter 7 First we give a short presentation of the concept and a version of Oseledets multiplicative ergodic theorem (MET). It is then shown that under certain integrability conditions for the Lévy measure Ruelle's integrability condition holds for solutions to the *SDE* under consideration and thus the MET provides the existence of a Lyapunov spectrum.
Chapter 8 is devoted to a formula of Furstenberg–Khasminskii type. It is shown that under some conditions the projection of the process to the sphere is a well defined and ergodic process. Then it is shown that the top Lyapunov exponent is generated by the ergodic theorem as an additive cocycle over the discrete time dynamics on the sphere. Finally a Furstenberg–Khasminskii type formula is obtained in the continuous time limit.

Part I

Variation of Jump diffusions

Chapter 1

Stochastic differential equations and stochastic flows

This introductory chapter shall provide the stochastic analysis used in this work. In Section 1.1 a concrete probability space, the Wiener–Poisson or *canonical* space, is defined and its relation with Lévy processes is discussed. Section 1.2 reviews the notion of semimartingales with spatial parameter and stochastic differential equations based on it. Under standard linear growth and Lipschitz conditions the existence and uniqueness of strong solutions and stochastic flows are presented.

1.1 Lévy processes and the Wiener–Poisson space

This first chapter discusses the fundamental building blocks for modelling (white) noise processes in continuous time – thus the objects on which all stochastic analysis, ergodic theorems and variational calculus in this thesis are based on. These objects are modelled on the *Wiener–Poisson* probability space, which is the canonical space to model continuous time random processes with stationary and independent (hence *i.i.d.*) increments. The chapter aims to review and fix terminology and standard results as well as to clarify notations. While the Wiener measure on the space of m-dimensional continuous functions $\mathscr{C}([0,T];\mathbb{R}^m)$ provides a standard m-dimensional *Brownian motion* as a continuous random input into the system, the Poisson space provides random discontinuities in the form of *jumps*. A generic noise process allowing for both a continuous part together with jumps with increments being stationary in time is known as a *Lévy process* and naturally modelled on the product space. The probability space explicitly given we will be able to investigate the sensitivity of the probability measure with respect to perturbations in Chapter 2 and 3.

1.1.1 The Wiener–Poisson probability space

In this introduction we fix a finite time horizon $T > 0$. Let Ω be the *canonical space* given as the product $\Omega = \Omega_W \times \Omega_N$ where

- Ω_W is the set of continuous functions $w \in \mathscr{C}_0([0,T];\mathbb{R}^m)$ starting at 0 at time $t=0$.
- Ω_N is the set of all *point clouds* (*configurations*) $\boldsymbol{u} \in \mathscr{U}([0,T];\mathbb{R}^m)$. Here
$$\mathscr{U}([0,T];\mathbb{R}^m) = \left\{\boldsymbol{u} = \{u_1, u_2, \ldots\} \text{ with } u_i \in (0,T] \times \mathbb{R}_0^m, i \in \mathbb{N}\right\} . \tag{1.1}$$
Given $\boldsymbol{u} \in \mathscr{U}([0,T];\mathbb{R}^m)$, every $u_i = (t_i, z_i)$ is interpreted as a jump of magnitude $z_i \in \mathbb{R}_0^m$ occurring at a jump-time $t_i > 0$. [1]

Any configuration $\boldsymbol{u} \in \mathscr{U}$ induces a *point process* $p_{\boldsymbol{u}}$ via the mapping
$$p_{\boldsymbol{u}}: \quad [0,T] \ni t \mapsto \begin{cases} z_i, & t = t_i \ , \\ 0, & \text{otherwise} \ . \end{cases}$$
Recall that a (random) function is called a *point process* if it takes values different from zero only on a countable subset of $(0,T]$ (*cf.* [IW89, Def.9.1]).
To give Ω an appropriate topology we endow both spaces $\mathscr{C}_0$ and $\mathscr{U}$ with the uniform norm $\|\cdot\|_\infty$ [2] and denote the corresponding product Borel σ-algebra by $\mathcal{F} = \mathcal{B}(\mathscr{C}_0) \otimes \mathcal{B}(\mathscr{U})$.
Define $\mathbb{P} = \mathbb{P}_W \otimes \mathbb{P}_N$ to be the product measure of

- the *Wiener measure* $\mathbb{P}_W$, i.e. a probability measure on Ω such that the *coordinate map*
$$\begin{aligned} W &: \Omega \to \mathscr{C}([0,T];\mathbb{R}^m) \ , \\ W(\omega)_t &= w_t \ , \ t \in [0,T] \end{aligned} \tag{1.2}$$
is a *Brownian motion* on $\mathbb{R}^m$,
- and a *Poisson measure* $\mathbb{P}_N$, i.e. a probability measure on Ω such that the counting measure
$$\begin{aligned} N &: \Omega \to \mathscr{M}([0,T] \times \mathbb{R}_0^m) \ , \\ N(\omega)(E) &= \#\{u_i : u_i \in E\} \ , \ E \in \mathcal{B}([0,T] \times \mathbb{R}_0^m) \end{aligned} \tag{1.3}$$
is a *Poisson random measure* on $[0,T] \times \mathbb{R}^m$ with a σ-finite intensity measure $\nu(\mathrm{d}z)\mathrm{d}t$, i.e.
$$\mathbb{E}\left[\ N(E)\ \right] = \int_E \nu(\mathrm{d}z)\mathrm{d}t \ , \ E \in \mathcal{B}([0,T] \times \mathbb{R}_0^m)$$
In particular the *compensated Poisson random measure* $\tilde{N}(\cdot) = N(\cdot) - \nu(\mathrm{d}z)\mathrm{d}t$ is a measure valued martingale in the sense that for every $E \in \mathcal{B}([0,T] \times \mathbb{R}_0^m)$ and $0 \leqslant s \leqslant t \leqslant T$
$$\mathbb{E}\left[\ \tilde{N}(E \cap [0,t] \times \mathbb{R}_0^m) | \mathscr{F}_s^N\ \right] = \tilde{N}(E \cap [0,s] \times \mathbb{R}_0^m) \ , \tag{1.4}$$
where $\mathscr{F}_s^N$ is the Borel σ-algebra $\mathcal{B}([0,s] \times \mathbb{R}_0^m)$. We also assume that under $\mathbb{P}_N$ the set
$$\{\boldsymbol{u} \in \mathscr{U} : \exists i \neq j \ , t_i = t_j, z_i \neq z_j\}$$
has measure zero. Hence there is only one jump at a time.

[1] Note that also finite point clouds are allowed.
[2] We stress that with the uniform topology $\mathscr{C}_0$ is separable but not $\mathscr{U}$ (see [Bil99]).

For further details we refer to the standard literature (e.g. [IW89, II.3]. We call $(\Omega, \mathbb{P})$ the *Wiener–Poisson* space.

1.1.2 Lévy processes

This section gives an overview on the class of Lévy processes which will be the source of randomness in our discussion. We do not wish to give a full characterization here and rather recall some important properties as well as make the reader familiar with terminology and notation. For further reading we refer to the extensive literature e.g. [Sat11, Sat14] and the monographs [App09, Ber96, Sat99].

Definition 1.1 (Lévy process)**.** A stochastic process $Z = (Z_t)_{t\in[0,T]}$ with values in $\mathbb{R}^m$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a LÉVY PROCESS if the following properties are fulfilled.

(i) It has *independent increments*, i.e for any $0 \leqslant t_0 < t_1 < \cdots < t_n \leqslant T,\ n \in \mathbb{N}$ the random variables $Z_{t_i} - Z_{t_{i-1}},\ i = 1, \dots, n$, are independent.

(ii) It is *time homogeneous*, i.e. the laws of $Z_{t+h} - Z_{s+h}$ do not depend on h for any s, t such that $s + h, t + h \in [0, T]$.

(iii) It is *continuous in probability*, i.e. for any t and $\varepsilon > 0$ we have $\lim_{h\to 0} \mathbb{P}(|Z_{t+h} - Z_t| > \varepsilon) = 0$.

For simplicity we will assume that $Z_0 = 0$. It can also be shown that such a process has a *càdlàg modification* which we will identify with Z. We thus assume that $\mathbb{P}$-almost all sample paths of Z are càdlàg (French acronym for *continu à droite, limites à gauche* meaning right continuous with left limits) (e.g. [App09, Chap. 2]).
It is clear from the independence and stationarity of increments that for any $t > 0$ the marginal distribution $\mu_t = \mathrm{Law}(Z_t)$ is *infinitely divisible*, which by definition says for any $n \in \mathbb{N}$ there exists a probability measure ρ on $\mathcal{B}(\mathbb{R}^m)$ such that $\mu_t = \rho * \cdots * \rho$ equals the n-fold convolution of ρ with itself. In fact it is easy to see that $\rho = \mu_{t/n}$.
In fact there is a one-to-one correspondence in the sense that any infinitely divisible distribution defines a Lévy process in law. The interplay between the two notions is extensively studied in the monograph [Sat99]. Any infinitely divisible distribution and thus the law of any Lévy process is determined by the celebrated *Lévy–Khintchine formula* which gives the Fourier transform or *characteristic function* of μ_t a specific form. In fact, there is a *drift vector* $b \in \mathbb{R}^m$, a symmetric positive semidefinite *(covariance-) matrix* $A \in \mathbb{R}^{m\times m}$ and a σ-finite *Lévy measure* ν on $\mathbb{R}^m_0$ satisfying

$$\int_{\mathbb{R}^m_0} |z|^2 \wedge 1\ \nu(\mathrm{d}z)\ , \tag{1.5}$$

such that the characteristic function of μ_t for any $t \geqslant 0$ with argument $x \in \mathbb{R}^m$ is given by

$$\mathbb{E}\left[\ \mathrm{e}^{\boldsymbol{i}\langle x, Z_t\rangle}\ \right] = \exp\left\{ t \left(\boldsymbol{i}\langle x, b\rangle + \langle x, Ax\rangle + \int_{\mathbb{R}_0^m} \mathrm{e}^{\boldsymbol{i}\langle x,z\rangle} - 1 - \boldsymbol{i}\langle x, z\rangle \mathbb{1}_{\{|z|\leqslant 1\}}\ \nu(\mathrm{d}z) \right) \right\} . \tag{1.6}$$

The triplet (b, A, ν) is commonly referred to as the *characteristic* or *Lévy triplet* of Z and in short we write $Z \sim (b, A, \nu)$. We remark that the cut-off at 1 in (1.6) is arbitrary and we could have taken any $\varepsilon \in (0, \infty)$. Of course one needs to adapt the drift vector b to

$$b_\varepsilon = b - \int_{\varepsilon < |z| \leqslant 1} z\nu(\mathrm{d}z) \ , \tag{1.7}$$

with the convention that $\int_{\varepsilon<|z|\leqslant 1} z\nu(\mathrm{d}z) = -\int_{1<|z|\leqslant\varepsilon} z\nu(\mathrm{d}z)$ whenever $\varepsilon > 1$. We have already lined out that the trajectories of Z are only continuous in probability so that in fact the mapping $t \mapsto Z_t \in \mathbb{R}^m$ may not be continuous. We denote by

$$\Delta Z_t = Z_t - Z_{t-} = Z_t - \lim_{s\nearrow t} Z_s \ , \tag{1.8}$$

the discontinuity or *"jump"* of Z at t which exists by Definition 1.1. A Lévy process comes naturally with a *counting measure* N defined as follows. For $0 \leqslant s < t < \infty$ and $E \in \mathcal{B}(\mathbb{R}_0^m)$ define

$$N\left(E \times (s,t]\right) = \#\left\{\Delta Z_r \in E\ ,\ r \in (s,t]\right\} \in \mathbb{N} \cup \{0, \infty\}\ , \tag{1.9}$$

the number of discontinuities located in E within the time interval $(s, t]$. As a consequence of Definition 1.1 N is finite whenever E is bounded away from zero. By (ii) N has further stationary increments in time. (i) implies that $N(\cdot)$ is independent when evaluated over disjoint sets. Hence the counting measure N is a Poisson random measure with intensity measure $\nu(\mathrm{d}z)\mathrm{d}t$. In particular

$$\tilde{N}(\mathrm{d}z\mathrm{d}t) = N(\mathrm{d}z\mathrm{d}t) - \nu(\mathrm{d}z)\mathrm{d}t\ , \tag{1.10}$$

is a martingale in t on every set of finite ν-measure. We also associate a *point process* to Z by assigning

$$p : t \mapsto \Delta Z_t\ ,\ t \in [0, T]\ . \tag{1.11}$$

Its discontinuities are identified with the realization of a point cloud $\boldsymbol{u} \in \mathscr{U}$ of the previous section.

With this terminology we can finally state the powerful characterization known as the *Lévy–Itô decomposition*. In fact every Lévy process Z can be represented by the decomposition

$$Z_t = bt + A^{\frac{1}{2}}W_t + \iint\limits_{\substack{|z|\leqslant 1\\ 0<s\leqslant t}} z\tilde{N}(\mathrm{d}z\mathrm{d}s) + \iint\limits_{\substack{|z|>1\\ 0<s\leqslant t}} zN(\mathrm{d}z\mathrm{d}s)\ , \tag{1.12}$$

where W is an m-dimensional Brownian motion independent of the Poisson random measure N and $A^{\frac{1}{2}}$ is any square root of the positive semidefinite matrix A in the Lévy triplet. It is now clear why the Wiener-Poisson space of the previous section is coined the "canonical" probability space for the construction of a Lévy process.

Remark 1.2. It is also worth noting that as soon as $\int_{|z|\leqslant 1} |z|\nu(\mathrm{d}z)$ is finite the jumps are summable and that there is no need to compensate the small jumps in (1.5).

We close this section with the observation that in what follows it is sufficient to consider the *filtration* $(\mathscr{F}_t)_{t\in[0,T]}$ on Ω generated by the Lévy process Z

$$\mathscr{F}_t = \bigcap_{\delta>0} \sigma\big(Z_s : 0 \leqslant s \leqslant t+\delta\big) \tag{1.13}$$

completed with the $\mathbb{P}$-null sets.

1.2 Semimartingales with spatial parameter, stochastic differential equations

1.2.1 Stochastic differential equations

For the stochastic analysis in this work we rely on the formalism of stochastic differential equations based on semimartingales with spatial parameters (cf. [Kun90, Kun04]). A stochastic differential equation (*SDE*) is given formally by an equation

$$\begin{cases} \mathrm{d}\xi_t = X(\xi_{t-}, \mathrm{d}t) \\ \xi_0 = x \in \mathbb{R}^d \end{cases}, \tag{1.14}$$

where X is a semimartingale with spatial parameter [3] X, i.e. a random field $\big(X(x,t) : x \in \mathbb{R}^d, t \geqslant 0\big)$ on the Wiener–Poisson space $(\Omega,\mathbb{P})$ of the form

$$X(x,t) = t\beta(x) + \sigma(x)W_t + \iint_{\mathbb{B}\times[0,t]} \gamma(x,z)\tilde{N}(\mathrm{d}z\mathrm{d}s) + \iint_{\mathbb{B}^c\times[0,t]} \gamma(x,z)N(\mathrm{d}z\mathrm{d}s)\ . \tag{1.15}$$

Here $\mathbb{B}$ is a (open) suitable ball around $0 \in \mathbb{R}^m$. The components β, σ and γ are Borel–measurable functions referred to as follows.

- $\beta : \mathbb{R}^d \to \mathbb{R}^d$ is the *drift coefficient*,
- $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ is called the *diffusion coefficient*,
- and $\gamma : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$ the *jump kernel*.

[3] The collection of processes $\big(X(x,\cdot)\big)_{x\in\mathbb{R}^d}$ forms a family of semimartingales in t indexed by $x \in \mathbb{R}^d$. We refer to the standard literature (e.g. [JS03] or [Pro04]).

If $\mathbb{B} = \mathbb{B}(\varepsilon) = \{z : |z| < \varepsilon\}$ is the ball of size $\varepsilon > 0$ we will sometimes use the notation

$$X_\varepsilon(x,t) = X_{\mathbb{B}}(x,t) = X(x,t) - \iint_{\mathbb{B}^c \times [0,t]} \gamma(x,z) N(\mathrm{d}z\mathrm{d}s) \ , \tag{1.16}$$

to refer to the semimartingale generator with bounded jumps.
We will impose standard linear growth – and Lipschitz conditions on the coefficients to ensure the existence and uniqueness of solutions to the stochastic differential equation (1.14).

Condition 1 (Linear growth, Lipschitz)**.** Assume that for some $p \geqslant 2$ there are positive functions K, L in $L^2 \cap L^p(\mathbb{R}^m, \nu)$ such that

$$\int_{\mathbb{B}} \left(K(z)^2 + K(z)^p + L(z)^2 + L(z)^p\right) \ \nu(\mathrm{d}z) < \infty \ , \tag{1.17}$$

and that for all $t \in [0,T], x, y \in \mathbb{R}^d$

$$\frac{|\beta(x)|}{1+|x|} \leqslant K(0) \ , \qquad |\beta(x) - \beta(y)| \leqslant L(0)|x-y|, \tag{1.18}$$

$$\frac{|\sigma(x)|}{1+|x|} \leqslant K(0) \ , \qquad |\sigma(x) - \sigma(y)| \leqslant L(0)|x-y|, \tag{1.19}$$

$$\frac{|\gamma(x,z)|}{1+|x|} \leqslant K(z) \ , \qquad |\gamma(x,z) - \gamma(y,z)| \leqslant L(z)|x-y| \ . \tag{1.20}$$

For $t > t_0 \geqslant 0$ consider now the following *SDE*

$$\begin{cases} \mathrm{d}\xi_t = X_{\mathbb{B}}(\xi_{t-}, \mathrm{d}t) \ , \\ \xi_{t_0} = \zeta \in \mathbb{R}^d \ , \end{cases} \tag{1.21}$$

where ζ is a $\mathscr{F}_0$-measurable random initial vector in $\mathbb{R}^d$. We have the following result ([Kun04, Sec.3.1., Sec.3.2., p.338ff.]).

Theorem 1.3 (existence and uniqueness I)**.** *Under the linear growth and Lipschitz condition 1 for some $p \geqslant 2$ (1.21) has a unique solution in L^p for any $\mathscr{F}_0$-measurable initial condition ζ in L^p . Furthermore for $\zeta = x \in \mathbb{R}^d$ deterministic, there exists a modification such that for every $t \geqslant 0$ the solution flow $\xi_t = \xi_t(x)$ is continuous with respect to the initial condition x in $\mathbb{R}^d$.*

Remark 1.4. Replacing X_ε in (1.21) by X still yields uniqueness and existence but the solution may not be in L^p.

Example 1.5 (Lévy process). Let $\beta(x) \equiv b \in \mathbb{R}^d$, $\sigma(x) \equiv A \in \mathbb{R}^{d \times m}$ positive semidefinite ($A \succcurlyeq 0$) and $\gamma(x,z) = z$. Then the solution ξ of (1.14) equals $\xi_t = x + Z_t$, where Z is a Lévy process with triplet (b, A, ν). Actually $X(x, \cdot) = Z$ for any $x \in \mathbb{R}^d$.

1.2.2 Lévy driven *SDE* and the Marcus canonical equation

The discussion so far introduced the concept of stochastic integration on the Wiener–Poisson space generated by spatially dependent semimartingales. While this concept treats the Wiener and the Poisson measure as two independent sources of randomness we now take the perspective of a Lévy processes as driver of the stochastic differential equation and thus a unified source of randomness.

Lévy driven *SDE*

We state the condition that will allow us to interpret the general *SDE* (1.14) as an *SDE* driven by a Lévy process.

Condition 2 (linearizability)**.** We assume that the jump kernel $\gamma : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$ is $\mathscr{C}^2(\mathbb{R}^m)$, i.e. twice continuously differentiable in the jump variable, and that

$$\partial_z \gamma(x,z)|_{z=0} = \gamma'(x) = \sigma(x), \ x \in \mathbb{R}^d \ .$$

The meaning of Condition 2 is obvious. The action of the Poisson measure linearized for small jump sizes is equivalent to the action of the Gaussian measure. In this case the semimartingale X driving equation (1.14) can be represented in terms of a Lévy process Z with characteristic $(0, \mathrm{Id}, \nu)$ by

$$X_{\mathbb{B}}(x,t) = t\tilde{\beta}(x) + \sigma(x) Z_t + \sum_{s \leqslant t} \tilde{\gamma}(x, \Delta Z_s) \ , \tag{1.22}$$

where the non-linear jump kernel $\tilde{\gamma}(x,z) = \gamma(x,z) - \gamma'(x)z$ satisfies

$$\int_{\mathbb{B}} |\tilde{\gamma}(x,z)| \ \nu(\mathrm{d}z) < \infty \ . \tag{1.23}$$

And the modified drift coefficient reads

$$\tilde{\beta}(x) = \beta(x) - \int_{\mathbb{B}} \tilde{\gamma}(x,z)\nu(\mathrm{d}z), \ \forall x \in \mathbb{R}^d \ . \tag{1.24}$$

Hence the generically nonlinear stochastic integral decomposes into a linear Itô integral of multiplicative type and a summable non-linear part. More generally we could define the integral against a generic Lévy process $Z \sim (b, A, \nu)$ and Lévy–Itô decomposition (1.12) by

$$\xi_t = x + \int_0^t (\tilde{\beta} + \sigma b)(\xi_s)\mathrm{d}s + \int_0^t \sigma(\xi_s) A^{\frac{1}{2}} \mathrm{d}W_s \tag{1.25}$$

$$+ \iint_{\mathbb{B}\times[0,t]} \sigma(\xi_{s-}) z \tilde{N}(\mathrm{d}z\mathrm{d}s) + \iint_{\mathbb{B}\times[0,t]} \tilde{\gamma}(\xi_{s-}, z) N(\mathrm{d}z\mathrm{d}s) \ , \tag{1.26}$$

where $\mathbb{B}$ is the unit ball.
A simple example of a linearizable jump kernel would be a jump kernel that is already linear. Therefore we make the following definition.

Definition 1.6 (*SDE* of multiplicative type)**.** If the jump kernel is of the form $\gamma(x,z) = \sigma(x)z$ we call the equation *SDE* OF MULTIPLICATIVE TYPE. We denote the equation also by

$$\xi_t = x + \int_0^t \sigma(\xi_{s-})\mathrm{d}Z_s \ . \tag{1.27}$$

Remark 1.7. Condition 2 is not a strong restriction on (1.14). Indeed, we could canonically separate the Wiener- and the Poisson random measure without affecting the law of ξ. We could define a new Lévy process $\hat{Z}$ on some $\mathbb{R}^{m_1+m_2}$ with $0 \leqslant m_1, m_2 \leqslant m$ such that W generates a Brownian motion on $\mathbb{R}^{m_1}$ and N a Poisson random measure on $\mathbb{R}^{m_2}$. Interpreting the jump kernel $\gamma : \mathbb{R}^d \times \mathbb{R}^{m_2} \to \mathbb{R}^d$ we would set

$$\hat{\sigma}(x) := \begin{pmatrix} \sigma(x) & \gamma'(x) \end{pmatrix} \in \mathbb{R}^{d\times(m_1+m_2)} \ , \tag{1.28}$$

and obtain a corresponding semimartingale with spatial parameter satisfying Condition 2.

Marcus' canonical equation

The prototype of an equation with such a linearizable jump kernel is the Marcus canonical equation [Mar81] which is a non-local generalization of Stratonovich *SDE* [KPP95]. [Fuj91, AK93]. We consider an m-dimensional Lévy process $Z \sim (b, A, \nu)$, where we may assume that $A = \mathrm{diag}(a_1^2, \dots, a_m^2), a_i \in \mathbb{R}$. Given m vector fields $\sigma_j : \mathbb{R}^d \to \mathbb{R}^d$ in $\mathscr{C}^2(\mathbb{R}^d;\mathbb{R}^d)$ we denote by $\sigma \in \mathbb{R}^{d\times m}$ the matrix with columns

$$\sigma_{\cdot j}(x) = \sigma_j(x) \ .$$

For $z \in \mathbb{R}^m$ we denote by $\phi^{\sigma z}$ the time one mapping of the solution flow $\phi : [0,1] \to \mathbb{R}^d$ of the *ODE*

$$\begin{cases} \dot{\phi} = \sum_{j=1}^m z^j \sigma_{\cdot j}(\phi) = \sigma(\phi)z \ , \\ \phi_0 = x \ , \end{cases} \tag{1.29}$$

i.e $\phi^{\sigma z}(x) = \phi(1)$. If we would consider a family $\phi^{\sigma z_i}(x), i = 1, \dots, n$ (for some $n \in \mathbb{N}$) with $z_i = W_{t\frac{i}{n}} - W_{t\frac{i-1}{n}}$ the increment of a Brownian motion, the concatenation $\psi_n := \phi^{\sigma z_n} \circ \dots \circ \phi^{\sigma z_1}$ would provide us with the so called *Wong–Zakai approximation* [WZ65] of the Stratonovich equation

$$\begin{cases} \mathrm{d}\xi_t = \sum_{j=1}^m \sigma_j(\xi_s) \circ \mathrm{d}W_s^j \ , \\ \xi_0 = x \ . \end{cases} \tag{1.30}$$

If we want to obtain a similar "local" behavior for the solution ξ of an *SDE* with jumps – in the sense that it respects the dynamics locally prescribed by the vector fields $\sigma_j, j =$

$1,\dots,m$ – it is natural to ask that the jumps of the solution $\Delta\xi_t$ are given by the solution curves to the ODE (1.29), i.e.

$$\Delta\xi_t = \phi^{\sigma\Delta Z_t} - \xi_{t-} \ . \tag{1.31}$$

To this aim we define a semimartingale with spatial parameter

$$\begin{aligned} X^\diamond(x,t) = \sigma(x)A^{\frac{1}{2}}W_t + \iint_{\mathbb{B}\times[0,t]} (\phi^{\sigma z}(x) - x)\,\tilde N(\mathrm{d}z\mathrm{d}t) \\ + t\left(\sigma(x)b + \frac{1}{2}\sum a_j^2 \nabla\sigma_{\cdot j}(x)\sigma_{\cdot j}(x) + \int_{\mathbb{B}} (\phi^{\sigma z}(x) - x - \sigma(x)z)\,\nu(\mathrm{d}z)\right) \ . \end{aligned} \tag{1.32}$$

We can now solve the *SDE*

$$\begin{cases} \mathrm{d}\xi_t = X^\diamond(\xi,\mathrm{d}t) \ , \\ \xi_0 = x \ . \end{cases} \tag{1.33}$$

Definition 1.8 (Marcus' canonical equation)**.** For vector fields $\sigma_j \in \mathscr{C}^2(\mathbb{R}^d;\mathbb{R}^d)$, with $j = 1,\dots,m$, the *SDE* (1.33) is called the MARCUS' CANONICAL EQUATION and is denoted by

$$\xi_t = x + \sum_{j=1}^m \int_0^t \sigma_j(\xi_{s-}) \diamond \mathrm{d}Z_s^j = x + \int_0^t \sigma(\xi_{s-}) \diamond \mathrm{d}Z_s \ . \tag{1.34}$$

Observe that if the Lévy triplet is given by $(0, \mathrm{Id}, 0)$ (no drift, no jumps), then the definition of (1.34) equals the definition of the Stratonovich equation (1.30). From (1.32) we have the following Itô–Stratonovich/Marcus correction.

Proposition 1.9 (Itô–Stratonovich correction)**.** *Denote by X the semimartingale generator of* (1.27)*, viz.*

$$X(x,t) = t\sigma(x)b + \sigma(x)A^{\frac{1}{2}}W_t + \iint_{\mathbb{B}\times[0,t]} (\sigma(x)z)\,\tilde N(\mathrm{d}z\mathrm{d}t) + \iint_{\mathbb{B}^c\times[0,t]} (\sigma(x)z)\,N(\mathrm{d}z\mathrm{d}t) \ .$$

We have the following correspondence between the semimartingale generators of X and $X^\diamond$ of (1.34)*, for any $x \in \mathbb{R}^m$ $t \in [0,T]$*

$$X^\diamond(x,t) = X(x,t) + t\frac{1}{2}\sum a_j^2\nabla\sigma_{\cdot j}(x)\sigma_{\cdot j}(x) + \iint_{\mathbb{R}^m\times[0,t]} (\phi^{\sigma z}(x) - x - \sigma(x)z)\,N(\mathrm{d}z\mathrm{d}s) \ .$$

Similarly to the case of Stratonovich equations this definition yields a *Leibniz rule* (see [FK85, AK93] or [KPP95]) as follows.

Proposition 1.10 (Leibniz rule)**.** *Let ξ be the unique strong solution to* (1.34) *and $F \in \mathscr{C}_b^2(\mathbb{R}^d)$. Then the following Leibniz rule holds true*

$$F(\xi_t) = F(x) + \int_0^t \nabla F(\xi_{s-})\sigma(\xi_{s-}) \diamond \mathrm{d}Z_s \ . \tag{1.35}$$

The "first order" calculus implied by the Leibniz rule has a variety of geometrical consequences that we will discuss briefly.

- *Coordinate free property*. We already mentioned the coordinate free property of solutions to (1.34) (*cf.* [Fuj91]). Suppose that the vector fields σ_j are all tangent to a submanifold $\mathbb{M} \subset \mathbb{R}^d$. If the initial value x is on $\mathbb{M}$ then the solution will be on $\mathbb{M}$ at any future time with probability one. We will verify this property in a simplified version in Chapter 6. Furthermore, the definition of the canonical equation does not depend on the choice of local coordinates on $\mathbb{M}$, and hence is "coordinate free". Therefore the definition can be extended to abstract manifolds (*cf.* [FK85, AK93]).

- *Stochastic flow of diffeomorphisms*. Stochastic differential equations with jumps generically do not give rise to stochastic flows of diffeomorphisms (see the discussion in Section 2.2.2). However solutions to the Marcus' canonical equations do generate stochastic flows of diffeomorphisms under fairly general conditions (e.g. [FK99]).

- *Wong–Zakai approximation.* We mentioned the approximation scheme of Wong and Zakai [WZ65] for Stratonovich equations (see also [IW89, Chap.IV, Thm.7.2, p.497]).

  The solution of (1.34) can be approximated in a similar manner by random *ODE*, driven by a suitable approximation of the Lévy process Z. Results in this direction are obtained in [KPP95] using a time integrated version of Z and in [Kun95] using a time dicretized version.

  These approximation schemes can also be used to obtain *support theorems of Stroock–Varadhan type.* (e.g. see [Kun99]).

Chapter 2

Derivatives and Variation

Throughout this chapter we consider an *SDE* on an possibly infinite time interval $[0,T]$

$$\begin{cases} \mathrm{d}\xi_t = X(\xi_{t-}, \mathrm{d}t) \ , \\ \xi_0 = x \in \mathbb{R}^d \ , \end{cases} \tag{2.1}$$

on the canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ defined in section 1.1.1 with a semimartingale generator

$$X(x,t) = t\beta(x) + \sigma(x)W_t + \iint_0^t \gamma(x,z)\tilde{N}(\mathrm{d}z\mathrm{d}s) \ , \ x \in \mathbb{R}^d, t \in [0,T] \ , \tag{2.2}$$

where the $\mathrm{d}z$ integral ranges over a suitable ball $\mathbb{B} \subset \mathbb{R}^m$ which is dropped from the notation. The coefficients β, σ, γ are vector- (resp. matrix-) functions that are Lipschitz continuous and of linear growth (*cf.* Condition 1) such that there exists a unique strong solution in $L^p(\Omega)$ for some $p \geqslant 2$.

2.1 L^p-spaces and L^p-derivatives

The notion of *(Fréchet-) derivative* we pursue in this thesis is essentially finite dimensional, albeit the state spaces are generally not. In this section we assume that the state space is a Banach space $(\mathscr{B}, |\cdot|_{\mathscr{B}})$ and a generic probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Following [BC86] we give the following definition.

Definition 2.1 (L^p-derivative)**.** Let $p \geqslant 1$ and Λ an open environment of $0 \in \mathbb{R}^n$, $n \in \mathbb{N}$. A family of random variables (functionals) $(F^\lambda)_{\lambda \in \Lambda}$ taking values in $\mathscr{B}$ indexed by $\Lambda \subset \mathbb{R}^n$ ($n \in \mathbb{N}$) is said to be L^p-DIFFERENTIABLE if $F^\lambda \in L^p(\Omega, \mathbb{P}; \mathscr{B})$ and the mapping

$$\Lambda \ni \lambda \mapsto F^\lambda \tag{2.3}$$

is Fréchet differentiable at zero as a mapping from $\mathbb{R}^d$ to $L^p(\Omega, \mathbb{P}; \mathscr{B})$. In other words there exists an L^p-DERIVATIVE $(\frac{\partial}{\partial\lambda}\big|_{\lambda=0} F^\lambda) \in L^p(\Omega, \mathbb{P}; \mathscr{B})$ satisfying

$$\|F^\lambda - F^0 - (\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} F^\lambda)\lambda\|_{L^p(\Omega;\mathscr{B})} = o(|\lambda|) \text{ as } \lambda \to 0 \ . \tag{2.4}$$

We observe that this definition immediately gives a product rule.

Lemma 2.2 (Product rule)**.** *Given two conjugate exponents $1 \leqslant p, q \leqslant \infty$ such that $1/p+1/q = 1$ (where $1+1/\infty = 1$) and two families $(F^\lambda)_{\lambda\in\Lambda}, (G^\lambda)_{\lambda\in\Lambda}$ with an L^p-derivative $\frac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda$ and an L^q-derivative $\frac{\partial}{\partial\lambda}\big|_{\lambda=0}G^\lambda$. Then product $F^\lambda G^\lambda$ has an L^1-derivative given by*

$$\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\big(F^\lambda G^\lambda\big) = \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda G^0 + F^0\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}G^\lambda \ . \tag{2.5}$$

Proof. With Hölder's inequality we immediately obtain for $\lambda \in \Lambda$

$$\begin{aligned}\mathbb{E}\big[\ &|F^\lambda G^\lambda - F^0G^0 - \big(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda G^0 + F^0\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}G^\lambda\big)\lambda\ \big] \\ &\leqslant \|F^\lambda\|_{L^p}\|G^\lambda - G^0 - \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}G^\lambda\lambda\|_{L^q} \\ &+ \|G^0\|_{L^q}\|F^\lambda - F^0 - \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda\lambda\|_{L^p} + \|F^\lambda - F^0\|_{L^p}\|\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}G^\lambda\lambda\|_{L^q}\ ,\end{aligned}$$

and the Lemma is proven. ∎

It follows that the set of random variables that are L^p-differentiable for any p is an algebra. Similarly we obtain a chain rule.

Proposition 2.3 (Chain rule)**.** *Let $\varphi : \mathscr{B} \to \mathbb{R}$ be twice Fréchet differentiable (in the usual sense) with derivatives of polynomial growth and $(F^\lambda)_{\lambda\in\Lambda}$ a family of L^p-differentiable random variables for any $p \geqslant 1$ with values in $\mathbb{R}^d$. Then the family $(\varphi(F^\lambda))_{\lambda\in\Lambda}$ is also L^p-differentiable for any $p \geqslant 1$ as a family of real valued random variables and its L^p-derivative is given by*

$$\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\varphi(F^\lambda) = \nabla\varphi(F^0)(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda)\ . \tag{2.6}$$

Proof. A first order Taylor expansion on Banach spaces (see e.g. [Zei86, Theorem 4.6A, p.148f]) for the twice Fréchet differentiable function φ gives that

$$\begin{aligned}\varphi(F^\lambda) - \varphi(F^0) &- \nabla\varphi(F^0)(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda)\lambda \\ &= \nabla\varphi(F^0)\,\big(F^\lambda - F^0 - (\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda)\lambda\big) + R\ ,\end{aligned}$$

where the together with the polynomial growth assumption for the second derivatives there is a $q \geqslant 1$ such that the remainder R satisfies

$$\begin{aligned}|R| &\leqslant \frac{1}{2}\sup_{\tau\in[0,1]}\|\nabla^2\varphi(F^0 + \tau(F^\lambda - F^0)\| \cdot |F^\lambda - F^0|^2_{\mathscr{B}} \\ &\lesssim \frac{1}{2}\left(1 + \sup_{\tau\in[0,1]}|F^0 + \tau(F^\lambda - F^0)|^q_{\mathscr{B}}\right) \cdot |F^\lambda - F^0|^2_{\mathscr{B}} \\ &\lesssim \big(1 + |F^0|^q_{\mathscr{B}} + |F^\lambda|^q_{\mathscr{B}}\big) \cdot |F^\lambda - F^0|^2_{\mathscr{B}}\end{aligned}$$

Similarly we may assume that $\|\nabla\varphi(F^0)\| \lesssim 1+|F^0|^q_{\mathscr{B}}$. By the Cauchy–Schwartz inequality we estimate

$$\begin{aligned}
&\|\varphi(F^\lambda) - \varphi(F^0) - \nabla\varphi(F^0)(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda)\lambda\|^p_{L^p} \\
&\quad\lesssim \mathbb{E}\big[\ \big(1+|F^0|^q_{\mathscr{B}}\big)^{2p}\ \big]^{\frac{1}{2}}\mathbb{E}\big[\ |F^\lambda - F^0 - (\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda)\lambda|^{2p}_{\mathscr{B}}\ \big]^{\frac{1}{2}} \\
&\quad\ + \mathbb{E}\big[\ \big(1+|F^\lambda|^q_{\mathscr{B}}+|F^0|^q_{\mathscr{B}}\big)^{2p}\ \big]^{\frac{1}{2}}\mathbb{E}\big[\ |F^\lambda - F^0|^{4p}_{\mathscr{B}}\ \big]^{\frac{1}{2}} \\
&\quad\lesssim \|F^\lambda - F^0 - (\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}F^\lambda)\lambda\|^p_{L^{2p}} + \|F^\lambda - F^0\|^p_{L^{4p}}\ ,
\end{aligned}$$

where the last inequality holds by a uniform estimate over Λ. It is easily seen that the last expression is of order $o(|\lambda|^p)$ by our assumptions. ∎

It will be necessary to consider the variation of matrix inverses. The following lemma is a consequence of the chain rule (2.6) and a version of [Wat84, Lemma 1, p.45] (see also [Nua06, Lemma 2.1.6, p.100] and [IW89, p.376]). We will not give a proof here.

Lemma 2.4. *Consider a family random real $m\times m$-square matrices $(A^\lambda)_{\lambda\in\Lambda}$ for some $m\in\mathbb{N}$. Suppose that on a set of full probability each $A^\lambda\in\mathbb{R}^{m\times m}$, $\lambda\in\Lambda$, is invertible and denote its inverse by $B^\lambda=(A^\lambda)^{-1}$. Assume further that $|\det A^0|^{-1}\in L^p(\Omega)$ and that for all indices $i,j\in\{1,\dots,m\}$ the family of entries $(A^\lambda_{ij})_{\lambda\in\Lambda}$ of A^λ is L^p-differentiable at 0 for every $p\geqslant 1$. Then the family $(B^\lambda)_{\lambda\in\Lambda}$ is (entry wise) L^p-differentiable at 0 and we have for all i,j the formula*

$$\frac{\partial}{\partial\lambda}\Big|_{\lambda=0}B^\lambda_{ij} = -\sum_{kl} B^0_{ik}B^0_{jl}\Big(\frac{\partial}{\partial\lambda}\Big|_{\lambda=0}A^\lambda_{lk}\Big)\ . \tag{2.7}$$

In matrix notation the formula reads

$$\frac{\partial}{\partial\lambda}\Big|_{\lambda=0}B^\lambda = -B^0\Big(\frac{\partial}{\partial\lambda}\Big|_{\lambda=0}A^\lambda\Big)B^0\ . \tag{2.8}$$

In our setting $\mathscr{B}$ will be either the Euclidean space $\mathbb{R}^d$ or the Banach space of càdlàg functions $\mathbb{D}([0,T];\mathbb{R}^d)$ equipped with the uniform norm[1].

2.2 Differentiability of stochastic flows

We established that the *SDE* (2.1) admits a unique strong solution for any initial condition $x\in\mathbb{R}^d$. We investigate properties of the *stochastic flow* mapping

$$\xi:\ (x,t)\mapsto\xi_t(x)\in L^p(\Omega,\mathbb{P};\mathbb{R}^d)\ . \tag{2.9}$$

The concept of L^p-derivatives comes therefore natural to investigate smoothness with respect to the initial condition $x\in\mathbb{R}^d$. Indeed, with $\Lambda=\mathbb{B}_1(\mathbb{R}^d)$, the open unit ball in $\mathbb{R}^d$, we may consider the family

$$F^\lambda = \xi_t(x+\lambda)\ ,\ \lambda\in\Lambda \tag{2.10}$$

of functionals either with values in $\mathbb{R}^d$ (for fixed x,t) or – as processes – with values in $\mathbb{D}=\mathbb{D}([0,T];\mathbb{R}^d)$ (with the uniform topology and fixed x and T).

[1]Note that $\mathbb{D}$ is complete but not separable with respect to the uniform norm (see [Bil99])

2.2.1 Differentiability – the Jacobian

To obtain any smoothness result we need to impose stricter conditions on the coefficients of (2.1). Indeed we rely on the following condition.

Condition 3 (uniform $\mathscr{C}^{1+\delta}$)**.** For $p \geqslant 2$ the coefficients of (1.14) are uniformly $\mathscr{C}^{1+\delta}$ for some $0 < \delta$ in the following sense. There are positive functions K' in $L^2 \cap L^{p+p\delta}(\mathbb{R}^m, \nu)$ and L' in $L^2 \cap L^p(\mathbb{R}^m, \nu)$, i.e.

$$\int \left(K'(z)^2 + K'(z)^{p+p\delta} + L'(z)^2 + L'(z)^p\right) \nu(\mathrm{d}z) < \infty \ , \tag{2.11}$$

such that for all $x, y \in \mathbb{R}^d$, $t \in [0, T]$ and $z \in \operatorname{supp}(\nu)$ we have

$$\|\nabla\beta(x,t)\| \leqslant K'(0) \ , \qquad \|\nabla\beta(x,t) - \nabla\beta(y,t)\| \leqslant L'(0)|x-y|^\delta \ , \tag{2.12}$$

$$\|\nabla\sigma(x,t)\| \leqslant K'(0) \ , \qquad \|\nabla\sigma(x,t) - \nabla\sigma(y,t)\| \leqslant L'(0)|x-y|^\delta \ , \tag{2.13}$$

$$\|\nabla\gamma(x,t,z)\| \leqslant K'(z) \ , \qquad \|\nabla\gamma(x,t,z) - \nabla\gamma(y,t,z)\| \leqslant L'(z)|x-y|^\delta \ . \tag{2.14}$$

We have the following ([Kun04, Thm.3.4., p.346]).

Theorem 2.5 (the Jacobian)**.** *Assume that the coefficients of X in* (2.2) *are all $\mathscr{C}^{1+\delta}$ with respect to x in the sense of Condition 3 for some $p \geqslant 2$. Then the $\mathbb{D}([0,T];\mathbb{R}^d)$ valued family $(F^\lambda)_{\lambda\in\Lambda}$ in* (2.10) *has an L^p-derivative at* 0 *and $x \in \mathbb{R}^d$ denoted by $\nabla\xi$. Furthermore the derivative process $\nabla\xi$ – in fact: the Jacobian of the stochastic flow – satisfies the matrix valued SDE*

$$\begin{cases} \mathrm{d}\nabla\xi_t = \nabla X(\xi_{t-}, \mathrm{d}t)\nabla\xi_{t-} \ , \\ \ \nabla\xi_0 = \mathrm{Id} \in \mathbb{R}^{d\times d} \ . \end{cases} \tag{2.15}$$

The matrix valued semimartingale generator ∇X in (2.15) is given by component-wise differentiation

$$\nabla X(x,t) = \nabla\beta(x)t + \nabla\sigma(x)W_t + \iint_0^t \nabla\gamma(x,z)\tilde{N}(\mathrm{d}z\mathrm{d}s) \ . \tag{2.16}$$

We do not give a proof here since it is basically analogous to the proof of Theorem 2.13 below. An alternative proof with the use of the extension theorem of Kolmogorov–Totoki is given in [Kun04, Thm. 3.4., p.346].

Remark 2.6. Note that under Condition 3 the *SDE* (2.15) is not Lipschitz and hence does not satisfy Condition 1. However the coupled *SDE* for ξ and $\nabla\xi$ considering (1.14) and (2.15) on $\mathbb{R}^d \times \mathbb{R}^{d\times d}$ has the obvious graded structure. Hence existence and uniqueness of a solution to (2.15) is guaranteed by Theorem B.4.

2.2.2 Flows of diffeomorphisms

Unlike in the Gaussian case the flow generated by (2.1) is in general not diffeomorphic. In fact, the jumps of the equation can easily destroy the injectivity of the flow. For example, if the Lévy measure ν associated to a Poisson random measure N is finite, the linear equation

$$\mathrm{d}\xi_t = B_0 \xi_{t-} \mathrm{d}t - \int \xi_{t-} N(\mathrm{d}z \mathrm{d}t) \tag{2.17}$$

with a matrix $B_0 \in \mathbb{R}^{d \times d}$ has a unique strong solution for any initial condition $x \in \mathbb{R}^d$. Furthermore the solution flow is differentiable and hence continuous by Theorem 2.5. The first jump of N occurs at an exponential random time $\tau > 0$ with rate $\nu(\mathbb{R}^m)$. However, at τ the solution jumps to the origin and will never leave due to linearity.
This simple example illustrates that in order for the solution to be homeomorphic it is necessary that the map $x \mapsto x + \gamma(x, z)$ should be one-to-one (see for instance [FK85] or [Kun04, Thm.3.11, p.356]).

Theorem 2.7. *Assume that Condition 3 holds for some $p \geqslant 2$ and that the maps $x \mapsto x+\gamma(x,z)$ are homeomorphic and its Jacobian matrix $[\mathrm{Id} + \nabla\gamma(x,z)]$ is invertible for ν-a.e. $z \in \mathbb{R}^m$. Then the SDE (2.1) generates a stochastic flow of diffeomorphisms in the sense that $\nabla\xi$ is invertible and in $L^p(\Omega, \mathbb{R}^{d\times d})$. Furthermore the inverse process $(\nabla\xi)^{-1}$ satisfies the SDE*

$$\begin{cases} \mathrm{d}(\nabla\xi)_t^{-1} = (\nabla\xi)_{t-}^{-1} Y(\xi_{t-}, \mathrm{d}t) \ , \\ \quad (\nabla\xi)_0^{-1} = \mathrm{Id} \in \mathbb{R}^{d\times d} \ , \end{cases} \tag{2.18}$$

with

$$Y(x,t) = -\nabla X(x,t) + t\nabla\sigma_j \nabla\sigma_j(x) + \iint_0^t \left[(\mathrm{Id} + \nabla\gamma)^{-1} (\nabla\gamma)^2\right](x,z)\ N(\mathrm{d}z\mathrm{d}s) \ . \tag{2.19}$$

Proof. For continuous semimartingales we refer to [Pro04, Theorem 48, p.326]. We only consider the case $\sigma \equiv 0$. First observe that

$$\left[(\mathrm{Id} + \nabla\gamma)^{-1}(\nabla\gamma)^2\right](x,z) = \left[\nabla\gamma - \mathrm{Id} + (\mathrm{Id} + \nabla\gamma)^{-1}\right](x,z) \ . \tag{2.20}$$

Then

$$\begin{aligned} Y(x,t) = t\left(-\nabla\beta(x) + \iint_0^t \left[(\mathrm{Id} + \nabla\gamma)^{-1}(\nabla\gamma)^2\right](x,z)\ \nu(\mathrm{d}z)\right) \\ + \iint_0^t \left[-\mathrm{Id} + (\mathrm{Id} + \nabla\gamma)^{-1}\right](x,z)\ \tilde{N}(\mathrm{d}z\mathrm{d}s) \ . \end{aligned} \tag{2.21}$$

Now let Φ be the unique solution to (2.18). Formally the jump at time t as a function of z satisfies

$$\left([\Phi\nabla\xi]_t - [\Phi\nabla\xi]_{t-}\right)(z) = \Phi_{t-}\left[(\mathrm{Id} + \nabla\gamma)^{-1}\right]\left[\mathrm{Id} + \nabla\gamma\right](\xi_{t-}, z)\nabla\xi_{t-} - [\Phi\nabla\xi]_{t-} = 0 \ .$$

Itô's formula[2] then yields

$$
\begin{aligned}
\mathrm{d}\,[\Phi\nabla\xi]_t &= \mathrm{d}t[\Phi\nabla\beta(\xi)\nabla\xi]_{t-} \\
&\quad + \mathrm{d}t\Phi_{t-}\left(-\nabla\beta(\xi_{t-}) + \int \left[(\mathrm{Id}+\nabla\gamma)^{-1}(\nabla\gamma)^2\right](\xi_{t-},z)\ \nu(\mathrm{d}z)\right)\nabla\xi_{t-} \\
&\quad + \int 0\ \tilde{N}(\mathrm{d}z\mathrm{d}t) \\
&\quad + \mathrm{d}t\int 0 - \Phi_{t-}\nabla\gamma(\xi_{t-},z)\nabla\xi_{t-} - \Phi_{t-}\left[-\mathrm{Id}+(\mathrm{Id}+\nabla\gamma)^{-1}\right](\xi_{t-},z)\nabla\xi_{t-}\ \nu(\mathrm{d}z)\ ,
\end{aligned}
$$

which is zero by (2.20). Hence $[\Phi\nabla\xi]_t \equiv \mathrm{Id}$ proving $\Phi_t = (\nabla\xi)_t^{-1}$. ■

Corollary 2.8. *Assume that the Lévy measure ν is supported on $\{z : \|\nabla\gamma(z)\| \leqslant c < 1\}$ for some constant c. Then the equation generates a flow of diffeomorphisms.*

2.3 Differentiation of the probability measure

This section elaborates the L^p-differentiability of the solution ξ to (2.1) with respect to perturbations of the probability measure. It is well known that in generic situations – if the noise in not additive – the law of ξ is singular with respect to linear perturbations

$$\omega \mapsto \omega + \theta \tag{2.22}$$

for any constant vector in $\theta \in \Omega$ (e.g. see [Bog10, Thm. 4.4.2 and Thm. 4.4.4]). It is therefore necessary to consider ω-dependent "random" perturbations. From a differential geometric perspective these correspond to "vector fields" in the language of [Bog10] that are in a certain sense tangent to the law of ξ.

It is plausible that in order to derive some smoothness of the solution to (2.1) with respect to a perturbation to the point process $\boldsymbol{u}$ we need to require a smooth dependence of the jump kernel γ with respect to z.

Condition 4. Assume that for some $p \geqslant 2$ the coefficients of the *SDE* satisfy Condition 3. Assume in addition that the diffusion coefficient σ and the partial derivative $\gamma' = \nabla_z\gamma$ are bounded and that γ' is Hölder continuous in z, i.e. there exists $\delta > 0$ and a constants $K'', L'' > 0$ such that

$$\|\sigma(x)\|, \|\gamma'(x,z)\| \leqslant K''\ , \quad \|\gamma'(x,z) - \gamma'(x,z')\| \leqslant L''|z - z'|^{\delta}\ , \tag{2.23}$$

for all $x \in \mathbb{R}^d$, and $z, z' \in \operatorname{supp}\nu$.

[2]See e.g. [IW89, Chap.II, Thm.5.1., p.66f]

2.3.1 A perturbation field

In order to study the absolute continuity of the image measure under d-dimensional functionals it is necessary to consider the perturbation with respect to d vector fields simultaneously. These vector fields can be thought of as a random basis of $\mathbb{R}^d$. To this aim we consider matrix valued perturbations θ of the Wiener–Poisson space $\Omega = \Omega_W \times \Omega_P$. We first introduce the set of simple perturbations. We rely on the notion of predictable[3] (matrix valued) processes, which form the class of generic integrands for stochastic integrals. We stress that the processes of interest will be deterministic functions in the random variables $\xi_{t-}, \nabla\xi_{t-}, (\nabla\xi)^{-1}_{t-}$, and as such, are predictable.

Condition 5 (Simple perturbation)**.** We denote the set of SIMPLE PERTURBATIONS by $\Theta_0 = \mathcal{H}_0 \times \mathcal{V}_0$ defined as a direct product where

- $\mathcal{H}_0$ is the collection of predictable matrix valued random fields $h : \Omega \times [0, T] \to \mathbb{R}^{m\times d}$ such that there exists a constant $0 < H < \infty$ such that

$$\|h\|_{[0,T]} = \sup_{0\leqslant t\leqslant T} \|h_t\| < H \ , \ \mathbb{P}\text{-a.s.} \ , \tag{2.24}$$

  i.e. h is bounded with respect to the operator matrix norm.

- $\mathcal{V}_0$ is the collection of predictable matrix valued random fields $v : \Omega \times [0, T] \times \mathbb{R}^m \to \mathbb{R}^{m\times d}$ continuously differentiable in $z \in \mathbb{R}^m$ such that there exists a deterministic continuous majorant function $V : \mathbb{R}^m \to [0, \infty)$ compactly supported in $\mathbb{R}^m \setminus \{0\}$, such that

$$\|v(\omega, t, z)\| + \|\nabla_z v(\omega, t, z)\| \leqslant V(z) \ , \ \forall t \in [0, T], \ \mathbb{P}\text{-a.s.} \ . \tag{2.25}$$

We remark that with this definition we have $V \in L^p(\mathbb{R}^m, \nu)$ for any Lévy measure ν on $\mathbb{R}^m$ and $p \geqslant 1$.
For a fixed $\theta = (h, v) \in \Theta_0$ and an environment Λ of $0 \in \mathbb{R}^d$ we define a family of transformations $(\mathcal{T}^\theta_\lambda)_{\lambda\in\Lambda}$ on Ω

$$\begin{gathered} \mathcal{T}^\theta_\lambda : \Omega \longrightarrow \Omega \\ \omega = (w, \boldsymbol{u}) \mapsto (w^{h\lambda}, \boldsymbol{u}^{v\lambda}) \end{gathered} \tag{2.26}$$

as follows.

- For $w \in \mathscr{C}([0, T]; \mathbb{R}^m)$ set

$$w^{h\lambda}_t(\omega) = w_t + \int_0^t h_s(\omega)\lambda \mathrm{d}s \ . \tag{2.27}$$

[3] For a definition we refer the reader e.g. to Definition 5.2 of [IW89, p.21]

- For $\boldsymbol{u} = \{(t_i, z_i)\ , i \in \mathbb{N}\} \in \mathscr{U}$ define a perturbed point cloud (configuration)

$$\begin{aligned} \boldsymbol{u}^{v\lambda}(\omega) &= \left\{(t_i, z^{v\lambda}(\omega, t_i, z_i))\ , i \in \mathbb{N}\right\}\ , \\ \text{where} \quad z^{v\lambda}(\omega, t, z) &= z + v(\omega, t, z)\lambda\ , \end{aligned} \tag{2.28}$$

and the associated perturbed counting measure given as the push-forward[4] of N under the transformation $\mathcal{T}^\theta_\lambda|_{\mathscr{U}}$ (restricted to $\mathscr{U}$)

$$N^{v\lambda}(\mathrm{d}z\mathrm{d}t) = (\mathcal{T}^\theta_\lambda|_{\mathscr{U}})_\# N(\mathrm{d}z\mathrm{d}t)\ . \tag{2.29}$$

Let us stress that only the locations of the jumps are perturbed. The jump times remain untouched (this allows us to work with the uniform topology on $\mathbb{D}$). We now investigate the effect of this transformation on the solution to the *SDE*.

Lemma 2.9. *Let ξ be the solution to the SDE* (2.1) *and $\theta \in \Theta_0$. Then for any $\lambda \in \Lambda$ the shifted process $\xi^\lambda := \xi \circ \mathcal{T}^\theta_\lambda$ satisfies under $\mathbb{P}$ the following SDE*

$$\begin{cases} \mathrm{d}\xi^\lambda_t = X_0(\xi^\lambda_{t-}, \mathrm{d}t) + \sigma(\xi^\lambda_{t-}) h_t \lambda \mathrm{d}t \\ \qquad + \displaystyle\int \gamma(\xi^\lambda_{t-}, z^{v\lambda}) \tilde{N}(\mathrm{d}z\mathrm{d}t) + \mathrm{d}t \int \left(\gamma(\xi^\lambda_{t-}, z^{v\lambda}) - \gamma(\xi^\lambda_{t-}, z)\right) \nu(\mathrm{d}z)\ , \\ \xi^\lambda_0 = x\ . \end{cases} \tag{2.30}$$

Proof. We focus on the jump kernel integral. Formally one obtains

$$(\mathcal{T}^\theta_\lambda)_\# \tilde{N}(\mathrm{d}z\mathrm{d}t) = N^{v\lambda}(\mathrm{d}z\mathrm{d}t) - \nu(\mathrm{d}z)\mathrm{d}t\ .$$

But the measure $\nu \times \mathrm{d}t$ is not the compensator of the shifted Poisson random measure $N^{v\lambda}$ under $\mathbb{P}$. We only demonstrate that the correction term in (2.30) is finite. For details we refer to [BGJ87, Lemma 6-20, p.66]. Indeed we have

$$|\int \gamma(\xi^\lambda_{t-}, z^{v\lambda}) - \gamma(\xi^\lambda_{t-}, z)\ \nu(\mathrm{d}z)| \leqslant \|\gamma'(\cdot,\cdot)\|_\infty \int |v(\omega, t, z)\lambda|\ \nu(\mathrm{d}z) < \infty\ .$$

■

2.3.2 Directional derivatives

We will establish the existence of strong derivatives in the direction of a larger class of perturbations than Θ_0.

Condition 6 (L^p-perturbation)**.** For $p \geqslant 2$ we define the set of L^p-PERTURBATIONS by the direct product $\Theta_p = \mathcal{H}_p \times \mathcal{V}_p$, where

[4] Recall that $(\mathcal{T}^\theta_\lambda|_{\mathscr{U}})_\# N(A) = N((\mathcal{T}^\theta_\lambda|_{\mathscr{U}})^{-1}(A))$ for all measurable sets A in $\mathscr{U}$.

- $\mathcal{H}_p$ is the collection of predictable matrix processes $h : \Omega \times [0,T] \to \mathbb{R}^{m\times d}$ with $\|h_t\|_{[0,T]} \in L^p(\Omega, \mathbb{P})$, i.e.
$$\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T} \|h_t\|^p \Big] < \infty \ .$$
- $\mathcal{V}_p$ is the collection of predictable matrix fields $v : \Omega \times [0,T] \times \mathbb{R}^m \to \mathbb{R}^{m\times d}$ such that $\|v\|_{[0,T]} \in L^1 \cap L^p(\Omega \times \mathbb{R}^m, \mathbb{P} \times \nu)$, i.e.
$$\mathbb{E}\Big[\int \Big\{ \sup_{0\leqslant t\leqslant T} \|v(\omega,t,z)\| + \sup_{0\leqslant t\leqslant T} \|v(\omega,t,z)\|^p \Big\} \nu(\mathrm{d}z) \Big] < \infty \ . \tag{2.31}$$

Remark 2.10. Obviously $\Theta_0 \subset \Theta_p$ for any $p \geqslant 2$. But in contrast to the set of simple perturbations Θ_0 the class of L^p-perturbations depends on the Lévy measure ν.

Example 2.11. In the case of additive noise we may choose v deterministically
$$v(\omega,t,z) = z^2\,\mathrm{e}^{-z} \cdot \mathrm{Id} \ . \tag{2.32}$$

Let us first prove the Lipschitz continuity of $\xi \circ \mathcal{T}^\theta_\lambda$ with respect to λ.

Lemma 2.12. *Assume that Condition 3 and Condition 4 hold with $p \geqslant 2$. We fix a perturbation $\theta = (h,v) \in \Theta_p$. Then for $\lambda \in \Lambda$*
$$\mathbb{E}\Big[\big|[\xi \circ \mathcal{T}^\theta_\lambda]_t - \xi_t\big|^p_{[0,T]} \Big] \lesssim_p |\lambda|^p \ . \tag{2.33}$$

Proof. Denote $\xi^\lambda = [\xi \circ \mathcal{T}^\theta_\lambda]$. We write down the dynamics of $\left[\xi^\lambda - \xi\right]$. With Lemma 2.9 we have
$$\begin{aligned} \mathrm{d}\left[\xi^\lambda - \xi\right]_t = \mathrm{d}t &\left\{\beta(\xi^\lambda) - \beta(\xi) + \sigma(\xi^\lambda)h\lambda\right\} \\ &+ \left\{\sigma(\xi^\lambda) - \sigma(\xi)\right\}\mathrm{d}W_t \\ &+ \int \left\{\gamma(\xi^\lambda, z^\lambda) - \gamma(\xi,z)\right\}\tilde{N}(\mathrm{d}z\mathrm{d}t) \\ &+ \mathrm{d}t \int \left\{\gamma(\xi^\lambda, z^\lambda) - \gamma(\xi^\lambda,z)\right\}\nu(\mathrm{d}z) \ . \end{aligned} \tag{2.34}$$

Using a first-order integral expansion we see that this equals
$$\begin{aligned} &= \mathrm{d}t \left(\int_0^1 \nabla\beta\left(\xi + \tau\left[\xi^\lambda - \xi\right]\right)\mathrm{d}\tau \left[\xi^\lambda - \xi\right] + \sigma\left(\xi^\lambda\right)h\lambda \right) \\ &+ \int_0^1 \nabla\sigma\left(\xi + \tau\left[\xi^\lambda - \xi\right]\right)\mathrm{d}\tau\left[\xi^\lambda - \xi\right]\mathrm{d}W_t \\ &+ \int\int_0^1 \Big\{ \nabla\gamma(\xi + \tau\left[\xi^\lambda - \xi\right], z)\left[\xi^\lambda - \xi\right] + \gamma'(\xi^\lambda, z + \tau\left[v\lambda\right])\left[v\lambda\right] \Big\}\mathrm{d}\tau\tilde{N}(\mathrm{d}z\mathrm{d}t) \\ &+ \mathrm{d}t \int\int_0^1 \Big\{ \nabla\gamma(\xi + \tau\left[\xi^\lambda - \xi\right], z)\left[\xi^\lambda - \xi\right] + \gamma'(\xi^\lambda, z + \tau\left[v\lambda\right])\left[v\lambda\right] \Big\}\mathrm{d}\tau\nu(\mathrm{d}z) \ , \end{aligned} \tag{2.35}$$

where in the last lines we expanded in the direction of $\big[\xi^\lambda - \xi\big]$ and $\big[z^\lambda - z\big] = [v\lambda]$ separately. We apply the L^p-estimates (Proposition B.2) bearing in mind the boundedness of the partial derivatives of the coefficients. We obtain

$$
\begin{aligned}
&\mathbb{E}\big[\ \big|\big[\xi^\lambda - \xi\big]\big|^p_{[0,t]}\ \big] \\
&\lesssim_p \mathbb{E}\big[\ \Big(\int_0^t K'(0)\big|\big[\xi^\lambda - \xi\big]_s\big| + K''\,|h_s\lambda|\,\mathrm{d}s\Big)^p\ \big] \\
&\quad + \mathbb{E}\big[\ \Big(\int_0^t K'(0)^2\big|\big[\xi^\lambda - \xi\big]_s\big|^2\mathrm{d}s\Big)^{p/2}\ \big] \\
&\quad + \mathbb{E}\big[\ \Big(\iint_0^t K'(z)^2\big|\big[\xi^\lambda - \xi\big]_s\big|^2 + K''^2|v\lambda|^2\nu(\mathrm{d}z)\mathrm{d}s\Big)^{p/2}\ \big] \\
&\quad + \mathbb{E}\big[\ \Big(\iint_0^t K'(z)^p\big|\big[\xi^\lambda - \xi\big]_s\big|^p + K''^p|v\lambda|^p\nu(\mathrm{d}z)\mathrm{d}s\Big)\ \big] \\
&\lesssim_p \Big(K'(0)^p + \int K'(z)^p\nu(\mathrm{d}z) + \big[\int K'(z)^2\nu(\mathrm{d}z)\big]^{p/2}\Big)\,\mathbb{E}\big[\ \int_0^t \big|\big[\xi^\lambda - \xi\big]\big|^p_{[0,s]}\,\mathrm{d}s\ \big] \\
&\quad + tK''^p\Big(\mathbb{E}\big[\ \|h\|^p_{[0,t]}\ \big] + \mathbb{E}\big[\ \Big(\int \|v\|^2_{[0,t]}\nu(\mathrm{d}z)\Big)^{p/2}\ \big] + \mathbb{E}\big[\ \int \|v\|^p_{[0,t]}\nu(\mathrm{d}z)\ \big]\Big)\,|\lambda|^p\ .
\end{aligned}
$$

Since the perturbations h, v satisfy Condition 6 we conclude with Gronwall's lemma that there exists a constant c_p such that

$$
\mathbb{E}\big[\ \big|\big[\xi^\lambda - \xi\big]\big|^p_{[0,t]}\ \big] \lesssim_p |\lambda|^p \exp\Big\{tc_p\Big(K'(0)^p + \int K'(z)^p\nu(\mathrm{d}z) + \big[\int K'(z)^2\nu(\mathrm{d}z)\big]^{p/2}\Big)\Big\}\ .
$$

■

From the above proof we see that for any $\lambda \in \Lambda \setminus \{0\}$ the difference quotient

$$
\widehat{\xi^\lambda_t} := \frac{1}{|\lambda|}\left([\xi \circ \mathcal{T}^\theta_\lambda]_t - \xi_t\right)\ , \tag{2.36}
$$

solves the *SDE* (2.34) multiplied by $\frac{1}{|\lambda|}$. It is possible to use the L^p-estimates of Proposition B.2 and the Kolmogorov–Totoki extension theorem to show that the random field $\{\widehat{\xi^{\varepsilon\lambda}_t} : x \in \mathbb{R}^d, \varepsilon > 0\}$ possesses an extension to $\varepsilon = 0$. Hence the process $\xi^{x,\lambda}$ indeed possesses an L^p-derivative $(\frac{\partial}{\partial\varepsilon}\big|_{\varepsilon=0}\xi^{\varepsilon\lambda})$ at $\lambda = 0$. It is the solution to the *SDE* obtained by sending ε to zero in (2.35). With $\bar\lambda = \lambda/|\lambda|$ being the unit vector in the direction of $\lambda \in \mathbb{R}^d$ $(\frac{\partial}{\partial\varepsilon}\big|_{\varepsilon=0}\xi^{\varepsilon\lambda})$ solves the d-dimensional *SDE*

$$
\begin{cases}
\mathrm{d}Y_t = \nabla X(\xi_{t-}, \mathrm{d}t)Y_{t-} + \sigma(\xi_{t-})h_t\bar\lambda\mathrm{d}t + \int \gamma'(\xi_{t-}, z)v(t,z)\bar\lambda N(\mathrm{d}z\mathrm{d}t)\ , \\
Y_0 = 0 \in \mathbb{R}^d\ .
\end{cases} \tag{2.37}
$$

Indeed we use a matrix valued version of this *SDE* to prove the existence of the L^p-derivative.

Theorem 2.13. *Assume that Condition 3 and Condition 4 hold with $p \geqslant 2$ and some $\delta > 0$. Then the process ξ_t possesses an L^p-derivative denoted by $\mathscr{D}_\theta \xi_t$ in the direction of $\theta \in \Theta_{p+p\delta}$ given as the unique solution to the matrix valued SDE*

$$\begin{cases} \mathrm{d}\mathscr{D}_\theta \xi_t = \nabla X(\xi_{t-}, \mathrm{d}t)\mathscr{D}_\theta \xi_{t-} + \sigma(\xi_{t-})h_t \mathrm{d}t + \displaystyle\int \gamma'(\xi_{t-}, z)v(t,z)N(\mathrm{d}z\mathrm{d}t) \ , \\ \mathscr{D}_\theta \xi_0 = 0 \in \mathbb{R}^{d\times d} \ . \end{cases} \tag{2.38}$$

Proof. To simplify the notation we drop indices referring to θ in the proof. Let $\mathscr{D}\xi_t$ denote the matrix whose columns are the collection $\{Y_t^{e_i}\}$ of solutions to (2.37) with $\bar{\lambda} = e_i$ for the standard basis $\{e_i\}$ of $\mathbb{R}^d$. By Theorem B.4 it is clear that $\mathscr{D}\xi$ exists and uniquely solves (2.38). We need to prove that

$$\mathbb{E}\big[\ |\xi^\lambda - \xi - \mathscr{D}\xi\lambda|^p\ \big] = o(|\lambda|^p) \ . \tag{2.39}$$

Consider the following evolution

$$\begin{aligned} \mathrm{d}\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right]_t = \mathrm{d}t &\left\{\beta(\xi^\lambda) - \beta(\xi) - \nabla\beta(\xi)\mathscr{D}\xi\lambda\right\} && (2.40)\\ &+ \left\{\sigma(\xi^\lambda) - \sigma(\xi) - \nabla\sigma(\xi)\mathscr{D}\xi\lambda\right\}\mathrm{d}W_t && (2.41)\\ &+ \mathrm{d}t\left\{\sigma(\xi^\lambda) - \sigma(\xi)\right\}h\lambda && (2.42)\\ &+ \int \left\{\gamma(\xi^\lambda, z^\lambda) - \gamma(\xi, z) - \nabla\gamma(\xi,z)\mathscr{D}\xi\lambda\right\}\tilde{N}(\mathrm{d}z\mathrm{d}t) && (2.43)\\ &+ \mathrm{d}t\int \left\{\gamma(\xi^\lambda, z^\lambda) - \gamma(\xi^\lambda, z)\right\}\nu(\mathrm{d}z) - \int \gamma'(\xi,z)v\lambda N(\mathrm{d}z\mathrm{d}t) \ , && (2.44) \end{aligned}$$

and observe that the last line equals

$$\mathrm{d}t\int \left\{\gamma(\xi^\lambda, z^\lambda) - \gamma(\xi^\lambda, z) - \gamma'(\xi,z)v\lambda\right\}\nu(\mathrm{d}z) - \int \gamma'(\xi,z)v\lambda\tilde{N}(\mathrm{d}z\mathrm{d}t) \ . \tag{2.45}$$

We can write the right hand side in line (2.40) as

$$\mathrm{d}t\left\{\int_0^1 \nabla\beta\left(\xi + \tau\left[\xi^\lambda - \xi\right]\right) - \nabla\beta(\xi)\mathrm{d}\tau\left[\xi^\lambda - \xi\right] + \nabla\beta(\xi)\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right]\right\} \ . \tag{2.46}$$

Analogously line (2.41) equals

$$\left\{\int_0^1 \nabla\sigma\left(\xi + \tau\left[\xi^\lambda - \xi\right]\right) - \nabla\sigma(\xi)\mathrm{d}\tau\left[\xi^\lambda - \xi\right] + \nabla\sigma(\xi)\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right]\right\}\mathrm{d}W_t \ , \tag{2.47}$$

and line (2.42) is just

$$\int_0^1 \nabla\sigma\left(\xi + \tau\left[\xi^\lambda - \xi\right]\right)\mathrm{d}\tau\left[\xi^\lambda - \xi\right]h\lambda \ . \tag{2.48}$$

We combine the martingale terms in (2.43) and (2.45) to deduce that the pure-jump martingale part is given by

$$\begin{aligned}
&\int \left\{\gamma(\xi^\lambda, z^\lambda) - \gamma(\xi, z) - \nabla\gamma(\xi, z)\mathscr{D}\xi\lambda - \gamma'(\xi, z)v\lambda\right\}\tilde{N}(\mathrm{d}z\mathrm{d}t) \\
&\quad = \int \Bigg\{ \int_0^1 \nabla\gamma\left(\xi + \tau\left[\xi^\lambda - \xi\right], z\right) - \nabla\gamma(\xi, z)\mathrm{d}\tau\left[\xi^\lambda - \xi\right] + \nabla\gamma(\xi, z)\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right] \\
&\qquad + \int_0^1 \gamma'\left(\xi^\lambda, z + \tau\left[v\lambda\right]\right) - \gamma'(\xi^\lambda, z)\mathrm{d}\tau\left[v\lambda\right] \Bigg\}\tilde{N}(\mathrm{d}z\mathrm{d}t) \ ,
\end{aligned}$$

where we expanded in the direction of $\left[\xi^\lambda - \xi\right]$ and $\left[z^\lambda - z\right] = [v\lambda]$ separately.
We hint that all terms can now be estimated in terms of $\left|\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right|$, $\left|\xi^\lambda - \xi\right|^{1+\delta}$, $|h\lambda|^{1+\delta}$ and $|v\lambda|^{1+\delta}$ by the boundedness and Lipschitz continuity of the partial derivatives of the coefficients.
We may now evoke the L^p-estimates of Proposition B.2 to obtain the bound

$$\begin{aligned}
&\mathbb{E}\Big[\ \left|\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right]\right|_{[0,t]}^p\ \Big] \\
&\lesssim_p \mathbb{E}\Big[\ \left(\int_0^t K'(0)\left|\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right]_s\right| + L'(0)\left|\left[\xi^\lambda - \xi\right]_s\right|^{1+\delta} + L'(0)\left|h_s\lambda\right|^{1+\delta}\mathrm{d}s\right)^p\ \Big] \\
&\quad + \mathbb{E}\Big[\ \left(\int_0^t K'(0)^2\left|\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right]_s\right|^2 + L'(0)^2\left|\left[\xi^\lambda - \xi\right]_s\right|^{2+2\delta}\mathrm{d}s\right)^{p/2}\ \Big] \\
&\quad + \mathbb{E}\Big[\ \Big(\iint_0^t K'(z)^2\left|\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right]_s\right|^2 + L'(z)^2\left|\left[\xi^\lambda - \xi\right]_s\right|^{2+2\delta} \\
&\qquad\quad + L''^2|v\lambda|^{2+2\delta}\nu(\mathrm{d}z)\mathrm{d}s\Big)^{p/2}\ \Big] \\
&\quad + \mathbb{E}\Big[\ \left(\iint_0^t K'(z)^p\left|\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right]_s\right|^p + L'(z)^p\left|\left[\xi^\lambda - \xi\right]_s\right|^{p+p\delta} + L''^p|v\lambda|^{p+p\delta}\nu(\mathrm{d}z)\mathrm{d}s\right)\ \Big] \\
&\lesssim_p \left(K'(0)^p + \int K'(z)^p\nu(\mathrm{d}z) + \left[\int K'(z)^2\nu(\mathrm{d}z)\right]^{p/2}\right)\mathbb{E}\Big[\ \int_0^t \left|\left[\xi^\lambda - \xi - \mathscr{D}\xi\lambda\right]\right|_{[0,s]}^p\mathrm{d}s\ \Big] \\
&\quad + \left(L'(0)^p + \int L'(z)^p\nu(\mathrm{d}z) + \left[\int L'(z)^2\nu(\mathrm{d}z)\right]^{p/2}\right)\mathbb{E}\Big[\ \int_0^t \left|\left[\xi^\lambda - \xi\right]\right|_{[0,s]}^{p+\delta p}\mathrm{d}s\ \Big] \\
&\quad + L'(0)^p\mathbb{E}\Big[\ \|h\|_{[0,t]}^{p+\delta p}\ \Big]|\lambda|^{p+p\delta} \\
&\quad + L''^p\left(\mathbb{E}\Big[\ \left(\int\|v\|_{[0,t]}^{2+2\delta}\nu(\mathrm{d}z)\right)^{p/2}\ \Big] + \mathbb{E}\Big[\ \int\|v\|_{[0,t]}^{p+p\delta}\nu(\mathrm{d}z)\ \Big]\right)|\lambda|^{p+p\delta}\ .
\end{aligned}$$

Lemma 2.12 assures that $\mathbb{E}\Big[\ \left|\left[\xi^\lambda - \xi\right]\right|_{[0,t]}^{p+\delta p}\ \Big] \lesssim_p |\lambda|^{p+\delta p}$ and (2.39) then follows with the aid of Gronwall's lemma. ■

2.3.3 An explicit formula for the derivative

Knowing that the L^p-derivative exists this section examines more closely its structure.

Theorem 2.14. *Assume that Condition 3 and Condition 4 hold with $p \geqslant 2$. Furthermore assume that the domain of integration $\mathbb{B}$ in (2.2) is contained in $\{z : \|\nabla\gamma(z)\| \leqslant c < 1\}$ such that ξ is diffeomorphic by Corollary 2.8. Then for any $\theta \in \Theta_p$ the L^p-derivative of ξ is given by the formula*

$$\mathscr{D}_\theta \xi_t = \nabla\xi_t \mathscr{A}_t^\theta \ , \tag{2.49}$$

where $\mathscr{A}_t^\theta$ is the random matrix given by the formula

$$\begin{aligned} \mathscr{A}_t^\theta &= \int_0^t \mathscr{A}^1(s)\mathrm{d}s + \iint_0^t \mathscr{A}^2(s,z)\ N(\mathrm{d}z\mathrm{d}s) \\ &= \int_0^t (\nabla\xi_{s-})^{-1}\sigma(\xi_{s-})h_s\mathrm{d}s \\ &\qquad + \iint_0^t (\nabla\xi_{s-})^{-1}(\mathrm{Id}+\nabla\gamma(\xi_{s-},z))^{-1}\gamma'(\xi_{s-},z)v(s,z)\ N(\mathrm{d}z\mathrm{d}s)\ , \end{aligned} \tag{2.50}$$

where $\gamma'(x,z) = \frac{\partial\gamma}{\partial z}(x,z)$.

Proof. Note that in general affine linear *SDE* as (2.38) can be solved by means of the variation-of-constants method (see [Jac82] or [BGJ87, pp.71f]). We give a proof for the case $\sigma, \beta, h = 0$. We also shorten the notation $\nabla\gamma = \nabla\gamma(\xi_{t-},z), v = v(t,z), \gamma' = \gamma'(\xi_{t-},z)$. Then

$$\begin{aligned} \mathrm{d}\nabla\xi_t &= \int \nabla\gamma\nabla\xi_{t-}\ \mathrm{d}\tilde{N}\ , \\ \mathrm{d}\mathscr{A}_t &= \int (\nabla\xi)_{t-}^{-1}(\mathrm{Id}+\nabla\gamma)^{-1}\gamma' v\ \mathrm{d}\tilde{N} \\ &\quad + \mathrm{d}t\int (\nabla\xi)^{-1}(\mathrm{Id}+\nabla\gamma)^{-1}\gamma' v\ \mathrm{d}\nu\ . \end{aligned}$$

Note that if there is a jump at time t we formally have

$$\begin{aligned} [\nabla\xi\mathscr{A}]_t - [\nabla\xi\mathscr{A}]_{t-} &= [(\mathrm{Id}+\nabla\gamma)\nabla\xi_{t-}]\left[\mathscr{A}_{t-} + (\nabla\xi)_{t-}^{-1}(\mathrm{Id}+\nabla\gamma)^{-1}\gamma' v\right] - [\nabla\xi\mathscr{A}]_{t-} \\ &= \nabla\gamma\,[\nabla\xi\mathscr{A}]_{t-} + \gamma' v\ . \end{aligned}$$

Hence by Itô's product rule we obtain

$$
\begin{aligned}
\mathrm{d}\,[\nabla\xi\mathscr{A}]_t = &\int \big(\nabla\gamma\,[\nabla\xi\mathscr{A}]_{t-} + \gamma' v\big)\ \mathrm{d}\tilde{N} \\
&+ \mathrm{d}t \int \big(\nabla\gamma\,[\nabla\xi\mathscr{A}]_{t-} + \gamma' v\big) - \nabla\xi_{t-}\ \big[(\nabla\xi)^{-1}(\mathrm{Id} + \nabla\gamma)^{-1}\gamma' v\ \big]_{t-} \\
&\qquad - [\nabla\gamma\nabla\xi]_{t-}\,\mathscr{A}_{t-}\ \mathrm{d}\nu \\
&+ \mathrm{d}t \int \nabla\xi_{t-}\ \big[(\nabla\xi)^{-1}(\mathrm{Id} + \nabla\gamma)^{-1}\gamma' v\ \big]_{t-}\ \mathrm{d}\nu \\
= &\int \nabla\gamma\,[\nabla\xi\mathscr{A}]_{t-}\ \mathrm{d}\tilde{N} + \int \gamma' v\ \mathrm{d}N\ ,
\end{aligned}
$$

which is exactly (2.38), in this specific case. ■

In honour of Paul Malliavin who put forward the stochastic calculus of variations we will call the matrix $\mathscr{D}_\theta\xi$ *Malliavin matrix* although the term refers strictly speaking to $\mathscr{D}_\theta\xi$ under a specific choice of θ made in Chapter 5. The matrix $\mathscr{A}_t^\theta$ is called the *reduced Malliavin matrix* (*cf.* [Mal97]). In fact one can consider $\mathscr{A}_t^\theta$ as the adapted part of the Malliavin matrix $\mathscr{D}_\theta\xi$ while $\nabla\xi_t$ is only $\mathcal{F}_t$ measurable (*cf.* [Hai08, p.21] and [Mal97, Rem. 5.3, p.244]).

2.4 Higher order derivatives

We recall a set of equations we obtained during the course of this based on a semimartingale generator X of the form

$$
X(x,t) = t\beta(x) + \sigma(x)W_t + \int \gamma(x,z)\tilde{N}(\mathrm{d}z\mathrm{d}t)\ , \quad x\in\mathbb{R}^d, t\in[0,T] \tag{2.51}
$$

as in (2.2). The solution ξ of the *SDE* (2.1), its Jacobian process $\nabla\xi$ and its L^p-derivative process follow the dynamics given by the equations (*cf.* (2.15), (2.38))

$$
\mathrm{d}\xi_t = X(\xi_{t-}, \mathrm{d}t)\ , \tag{2.52}
$$
$$
\mathrm{d}\nabla\xi_t = \nabla X(\xi_{t-}, \mathrm{d}t)\nabla\xi_{t-}\ , \tag{2.53}
$$
$$
\mathrm{d}\mathscr{D}_\theta\xi_t = \nabla X(\xi_{t-}, \mathrm{d}t)\mathscr{D}_\theta\xi_{t-} + \sigma(\xi_{t-})h_t\mathrm{d}t + \int \gamma'(\xi_{t-}, z)v(t,z)N(\mathrm{d}z\mathrm{d}t)\ . \tag{2.54}
$$

The strategy is to combine all these equations into a single high dimensional equation and to show that the solution is L^p-differentiable. Indeed one can prove the following.

Theorem 2.15. *Assume that for some $k\in\mathbb{N}$ and $\delta\in(0,1)$ the coefficients of X are all $\mathscr{C}^{k+\delta}$ in x and z with all partial derivatives bounded. Then ξ possesses L^p-derivatives ($p\geqslant 2$) in the direction of $\theta\in\Theta_p$ up to order k.*

We do not give a full proof since it is lengthy and technical and instead refer to [BGJ87]. We will sketch the arguments.

Sketch of Proof. The proof of Theorem 2.13 exploits the lower triangular Lipschitz property of the graded structure that also guarantees the existence of a solution to the combined equation. Indeed we have the grading $x = (x^1, x^2, x^3) \in \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times d}$ of the form

$$\widetilde{X}(x,t) = \begin{pmatrix} X(x^1,t) \\ \nabla X(x^1,t)x^2 \\ \nabla X(x^1,t)x^3 + t\sigma(x^1)h_t + \iint_0^t \gamma'(x^1,z)v(s,z)N(\mathrm{d}z\mathrm{d}s) \end{pmatrix} . \tag{2.55}$$

Under the assumptions and by Theorem B.4 there exists a solution to the *SDE*

$$\begin{cases} \mathrm{d}\Phi_t = \widetilde{X}(\Phi_{t-}, \mathrm{d}t) \\ \Phi_0 = (x, \mathrm{Id}, \mathbf{0})^* \end{cases} .$$

Similarly to Theorem 2.13 we may find that Φ has an L^p derivative in the direction of $\theta \in \Theta_0$ that solves

$$\begin{cases} \mathrm{d}\mathscr{D}_\theta\Phi_t = \nabla\widetilde{X}(\Phi_{t-}, \mathrm{d}t) + \widetilde{\sigma}(\Phi_{t-})h_t\mathrm{d}t + \int \widetilde{\gamma}'(\Phi_{t-}, z)v(t_z)N(\mathrm{d}z\mathrm{d}t) , \\ \mathscr{D}_\theta\Phi_0 = \mathbf{0} \in \mathbb{R}^{d(d+2d^2)} , \end{cases} \tag{2.56}$$

where $\widetilde{\sigma}, \widetilde{\gamma}'$ are the diffusion coefficient matrix and the derivative with respect to z of the jump kernel of $\widetilde{X}$. We then identify the components

$$\Phi_t = (\mathscr{D}_\theta\xi_t, \mathscr{D}_\theta\nabla\xi_t, \mathscr{D}_\theta^2\xi_t) \in \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times d\times d} \times \mathbb{R}^{d\times d\times d} .$$

This procedure can now be iterated up to order k. ■

Chapter 3

The integration-by-parts formula

In functional analysis the concept of *weak derivatives* extends the notion of a derivative from a "nice" space, say, of smooth functions via the integration-by-parts formula to its completion with respect to a certain Sobolev norm.
Here we follow a similar approach due to Bismut [Bis81]. We will derive an *integration-by-parts* formula for the L^p-derivative described in (2.38) for "nice" perturbations $\theta \in \Theta_0$. We then extend the formula to a suitable completion Θ of Θ_0.

3.1 Derivation of integration-by-parts

3.1.1 Bismut's idea of deriving integration-by-parts. An abstract motivation

Bismut [Bis81] and earlier [Pit63, Pit64] considered shifts on the path space $\Omega = \mathscr{C}([0,T])$ or $\Omega = \mathbb{D}([0,T])$ that leave the law of a given functional, e.g. the solution to the *SDE* (2.1) invariant. Assume that we have a measurable group of transformations $(\mathcal{T}_\lambda)_{\lambda\in\mathbb{R}}$ on Ω. The group induces a group of transformations on bounded measurable functionals $\varphi \in \mathscr{B}_b(\Omega)$ by assigning

$$\mathcal{T}_\lambda\varphi(\omega) = \varphi(\mathcal{T}_\lambda(\omega)) \ . \tag{3.1}$$

This group then has an adjoint action on measures $\mu \in \mathscr{M}(\Omega, \mathcal{A})$ where the image measure $(\mathcal{T}_\lambda)_\#\mu$ (the push forward) is characterized by

$$\int_\Omega \varphi(\omega)(\mathcal{T}_\lambda)_\#\mu(\mathrm{d}\omega) := \int_\Omega \varphi(\mathcal{T}_\lambda\omega)\mu(\mathrm{d}\omega) \ , \quad \forall\varphi \in \mathscr{B}_b(\Omega) \ . \tag{3.2}$$

Assume now that for a fixed μ the push forward measures are absolutely continuous $(\mathcal{T}_\lambda)_\#\mu \ll \mu$ for $|\lambda|$ small enough. This property is referred to as *ray continuity* of the measure μ.
Assume further that the family of Radon–Nikoým densities $(\varrho_\lambda)_{\lambda\in\mathbb{R}}$ has a derivative at $\lambda = 0$ with $\partial_\lambda\varrho_\lambda|_{\lambda=0} \in L^1(\mathrm{d}\mu)$.

For any functional $\varphi \in \mathscr{B}_b(\Omega)$ we have

$$\int_\Omega \varphi \mathrm{d}\mu = \int_\Omega \varphi \circ (\mathcal{T}_\lambda)^{-1} \circ \mathcal{T}_\lambda \mathrm{d}\mu = \int_\Omega \varphi \circ (\mathcal{T}_\lambda)^{-1} \mathrm{d}(\mathcal{T}_\lambda)_\# \mu = \int_\Omega (\mathcal{T}_{-\lambda}\varphi) \varrho_\lambda \mathrm{d}\mu \ .$$

Now since $\mathcal{T}_{-\lambda}\varphi$ is bounded and ϱ_λ differentiable in L^1 we obtain an integration-by-parts formula differentiating both sides at $\lambda = 0$

$$\int_\Omega \partial_\theta \varphi \mathrm{d}\mu = \int_\Omega \varphi \partial_\lambda \varrho_\lambda|_{\lambda=0} \mathrm{d}\mu \ . \tag{3.3}$$

We may define a weak derivative of the measure μ in the direction of $\mathcal{T}_\lambda$ as the signed measure given by

$$\partial\mu := -\partial_\lambda \varrho_\lambda|_{\lambda=0} \mathrm{d}\mu \ . \tag{3.4}$$

This notion of a derivative of a measure is due to S.V. Fomin (see also [Bog10, Def. 3.1.6]).

3.1.2 A Girsanov density

From now on we continue to work with the solution to the *SDE* (2.1) and the conditions imposed in Chapter 2. To obtain an integration-by-parts formula for the jump measure we need additional smoothness.

Condition 7. The Lévy measure ν admits a density $\varrho \in \mathscr{C}^{1+\delta}(\mathbb{R}^{m'} \setminus \{0\})$ with respect to the Lebesgue measure on some $\mathbb{R}^{m'} \subset \mathbb{R}^m$ for some $m' \leqslant m$. This density is strictly positive in an environment of $0 \in \mathbb{R}^{m'}$ (may be infinite at $\{0\}$ itself).

We usually consider ϱ as function on $\mathbb{R}^m$ not depending on the additional coordinates. No further notational distinction between the density and the Lebesgue measure on $\mathbb{R}^m$ and $\mathbb{R}^{m'}$ is made.

Remark 3.1. The condition that the Lévy density ϱ is positive in an environment of 0 and therefore has mass in all directions of $\mathbb{R}^{m'}$ is a simplification of this thesis. Thinking one-dimensional smoothing could also be obtained by only positive jumps, without any negative jumps. Indeed, the correct condition is the so called *sector condition* (see [Nor88]) that we will not state here. Morally it requires that the Lévy density ϱ is positive in the intersection of an environment of zero with the interior of a cone in $\mathbb{R}^{m'}$.

Let us fix a perturbation $\theta \in \Theta_0$ and recall the definition of the transformations $\mathcal{T}^\theta_\lambda$ in (2.26) for $\lambda \in \Lambda$. We will show that the push forward measure $\mathbb{P}^\lambda := (\mathcal{T}^\theta_\lambda)_\# \mathbb{P} = \mathbb{P} \circ (\mathcal{T}^\theta_\lambda)^{-1}$ is absolutely continuous with respect to $\mathbb{P}$. In what follows we construct the Radon–Nikodým density of $\mathbb{P}^\lambda$ with respect to $\mathbb{P}$ by an exponential martingale process $\mathcal{Z}^\lambda$. Observe that the functional determinant of the Jacobian of the transformation $z \mapsto z^{v\lambda} = z + v(\omega, t, z)\lambda$ for each ω, t, z and λ is given by

$$\det(\mathrm{Id} + \sum_{i \leqslant d} \nabla_z v^{\cdot i}(\omega, t, z)\lambda^i) = \det(\mathrm{Id} + \nabla_z v(\omega, t, z)\lambda) \ . \tag{3.5}$$

Denote by $\varrho^{v\lambda}(\omega,t,z) = \varrho(z+v(\omega,t,z)\lambda)$ the shifted Lévy density. We define a jump functional determinant for each ω, t, z and λ by

$$\begin{aligned} J^{v\lambda} = J^{v\lambda}(\omega,t,z) &= \det(\mathrm{Id} + \sum_{i\leqslant d} \nabla_z v^{\cdot i}(\omega,t,z)\lambda^i) \frac{\varrho^{v\lambda}(\omega,t,z)}{\varrho(z)} \\ &= \big[\det(\mathrm{Id} + \nabla_z v\lambda)\frac{\varrho^{v\lambda}}{\varrho}\big](\omega,t,z)\ , \end{aligned} \tag{3.6}$$

with the convention that $J^{v\lambda}(\omega,t,z) = 1$ if $\varrho(z) = 0$. This functional determinant has an L^p-derivative at $\lambda = 0$.

Lemma 3.2. *There is a deterministic function V' of compact support in $\mathbb{R}^m \setminus \{0\}$ such that $\|\frac{\partial}{\partial\lambda} J^{v\lambda}(\omega,t,z)\| < V'(z)$ for all $t \in [0,T]$ almost surely. (This implies $V' \in L^p(\nu)$ for all $p > 0$.) At $\lambda = 0$ we have the expressions*

$$\tfrac{\partial}{\partial\lambda} J^{v\lambda}\big|_{\lambda=0} = \mathbf{div}_z\, v + \frac{\nabla \varrho v}{\varrho} = \mathbf{div}_z\, v + v\nabla\log(\varrho) = \varrho^{-1}\,\mathbf{div}_z(v\varrho)\ . \tag{3.7}$$

Proof. We may assume that ϱ is positive in an environment of $z \in \mathbb{R}^{m'}$. Jacobi's formula states that $\frac{\mathrm{d}}{\mathrm{d}t}\det A(t) = \det A(t)\,\mathrm{Tr}(A^{-1}\frac{\mathrm{d}}{\mathrm{d}t}A(t))$ if a function A maps t into the invertible matrices. Since $|\nabla v| < V$ bounded we have for sufficiently small $|\lambda|$

$$\tfrac{\partial}{\partial\lambda_j} J^{v\lambda} = \det(\mathrm{Id} + \nabla v\lambda)\Big(\mathrm{Tr}\big((\mathrm{Id} + \nabla v\lambda)^{-1}\nabla v e_j\big) + \frac{\nabla\varrho(\cdot + v\lambda)ve_j}{\varrho}\Big) \tag{3.8}$$

with $e_1,\dots,e_m$ the canonical basis of $\mathbb{R}^m$. At $\lambda = 0$ this implies (3.7). Since ∇v is bounded we estimate

$$\begin{aligned} \|\tfrac{\partial}{\partial\lambda} J^{v\lambda}\| &\lesssim (1+|v\lambda|)\left\{\|\nabla v\| + \Big(\frac{|v\lambda|^\delta}{\varrho} + \frac{|\nabla\varrho|}{\varrho}\Big)\|v\|\right\} \\ &\leqslant (1+V|\lambda|)\left\{V + \frac{V^{1+\delta}}{\varrho}|\lambda|^\delta + \frac{V|\nabla\varrho|}{\varrho}\right\} \end{aligned} \tag{3.9}$$

If we recall that V is compactly supported in $\mathbb{R}^m \setminus \{0\}$ it is easy to see that the right hand side is in $L^p(\mathbb{R}^m,\nu)$ under Condition 7. Up to the constants we omitted we may define V' to be this right hand side. ■

Corollary 3.3. *It follows immediately that for any $p > 0$ we have*

$$|J^{v\lambda} - 1| \leqslant |\lambda| V'(z) \in L^p(\mathbb{R}^m,\nu)\ . \tag{3.10}$$

For every $\lambda \in \Lambda$ we define a real valued true martingale given by

$$M_t^{v\lambda} = -\int_0^t (h_s\lambda)^*\mathrm{d}W_s + \iint_0^t (J^{v\lambda}(s,z) - 1)\tilde{N}(\mathrm{d}z\mathrm{d}s)\ , \quad t \in [0,T]\ . \tag{3.11}$$

The corresponding exponential martingale (its Doléans–Dade exponential) $\mathcal{Z}^\lambda = \mathcal{E}(M^{v\lambda})$ (e.g. [Pro04]) is the solution to the affine equation

$$\mathcal{Z}_t^\lambda = 1 - \int_0^t \mathcal{Z}_{s-}^\lambda (h_s\lambda)^*\mathrm{d}W_s + \iint_0^t \mathcal{Z}_{s-}^\lambda (J^{v\lambda}(s,z) - 1)\tilde{N}(\mathrm{d}z\mathrm{d}s)\ . \tag{3.12}$$

Lemma 3.4. *For any* $\lambda \in \Lambda$ *and* $\theta = (h, v) \in \Theta_0$ *the solution is given explicitly and reveals the product structure* $\mathcal{Z}_t^\lambda = \mathcal{Z}_t^{h\lambda} \times \mathcal{Z}_t^{v\lambda}, t \in [0, T]$ *with*

$$\mathcal{Z}_t^{h\lambda} = \exp\left\{ -\int_0^t (h_s\lambda)^* \mathrm{d}W_s - \frac{1}{2}\int_0^t |h_s\lambda|^2 \mathrm{d}s \right\} \ , \tag{3.13}$$

$$\begin{aligned} \mathcal{Z}_t^{v\lambda} = \exp\Bigg\{ &\iint_0^t \log(J^{v\lambda}(s,z)) \ \tilde{N}(\mathrm{d}z\mathrm{d}s) \\ &- \iint_0^t (J^{v\lambda}(s,z) - 1 - \log(J^{v\lambda}(s,z))\nu(\mathrm{d}z)\mathrm{d}s \Bigg\} \ . \end{aligned} \tag{3.14}$$

Proof. For simplicity we only prove the non Gaussian case with $h \equiv 0$. Applying Itô's formula to the exponential function $x \mapsto \mathrm{e}^x$ we obtain by straightforward calculation

$$\begin{aligned} \mathrm{d}\mathcal{Z}_t^{v\lambda} &= \int \mathcal{Z}_{t-}^{v\lambda} \left(\exp\left\{\log(J^{v\lambda}(t,z))\right\} - 1\right) \ \tilde{N}(\mathrm{d}z\mathrm{d}t) \\ &\quad + \int \mathcal{Z}_{t-}^{v\lambda} \left(\exp\left\{\log(J^{v\lambda}(t,z))\right\} - 1 - \log(J^{v\lambda}(t,z))\right) \nu(\mathrm{d}z) \\ &\quad - \int \mathcal{Z}_{t-}^{v\lambda} \left(J^{v\lambda}(t,z) - 1 - \log(J^{v\lambda}(t,z))\right) \nu(\mathrm{d}z)\mathrm{d}t \\ &= \int \mathcal{Z}_{t-}^{v\lambda} \left(J^{v\lambda}(t,z) - 1\right) \ \tilde{N}(\mathrm{d}z\mathrm{d}t) \ . \end{aligned}$$

Which concludes the proof. ■

This product structure allows to interpret $\mathcal{Z}^\lambda$ as a Girsanov density for a product measure on Wiener–Poisson space with respect to $\mathbb{P}$. In fact we can now define the probability measure $\mathbb{P}^\lambda$ on Ω via $\frac{\mathrm{d}\mathbb{P}^\lambda}{\mathrm{d}\mathbb{P}} = \mathcal{Z}^\lambda$ and denote the corresponding expectation by $\mathbb{E}^\lambda$. Under this measure the process $\xi_t^\lambda = \xi_t \circ \mathcal{T}_\lambda^\theta$ has the same law as ξ under $\mathbb{P}$.

Proposition 3.5. *We have* $\mathbb{P}^\lambda \circ (\mathcal{T}_\lambda^\theta)^{-1} = \mathbb{P}$. *In particular we have for* $\Psi \in L^1(\mathbb{P})$

$$\mathbb{E}^\lambda[\Psi \circ \mathcal{T}_\lambda^\theta] = \mathbb{E}\big[\ (\Psi \circ \theta^\lambda)\mathcal{Z}^\lambda \ \big] = \mathbb{E}\Psi \ . \tag{3.15}$$

Proof. To prove the claim we need to show that under $\mathbb{P}^\lambda$, $W \circ \mathcal{T}_\lambda^\theta$ is a Brownian motion and $N \circ \mathcal{T}_\lambda^\theta$ is a Poisson random measure with intensity measure ν. We have already argued that the product structure of $\mathcal{Z}^\lambda = \mathcal{Z}^{h\lambda} \times \mathcal{Z}^{v\lambda}$ accounts for a change of the Wiener measure $\mathbb{P}_W$ on Ω_W and of $\mathbb{P}_N$ on Ω_N. For the Wiener measure we quote the Girsanov theorem (e.g. [Pro04, Theorem 46, p.143]). It assures that

$$[W \circ \mathcal{T}_\lambda^\theta]_t = W_t + \int_0^t h_s\lambda \mathrm{d}s$$

is a Brownian motion under $\mathbb{P}^{h\lambda}$ defined by $\frac{\mathrm{d}\mathbb{P}^{h\lambda}}{\mathrm{d}\mathbb{P}} = \mathcal{Z}^{h\lambda}$. To prove that $N \circ \mathcal{T}_\lambda^\theta$ is a Poisson random measure with intensity measure ν we use the characterization of random measures via Laplace functionals (e.g. [Çin11, Prop.1.4, p.244]). The Laplace functional acts on positive measurable functions $\phi \in L^1(\mathbb{R}^m \times [0,T], \nu \times \mathrm{d}t)$ by

$$\mathbb{E}^\lambda\big[\,\mathrm{e}^{-\iint_0^t \phi(s,z)[N\circ\mathcal{T}_\lambda^\theta](\mathrm{d}z\mathrm{d}s)}\big] = \mathbb{E}\big[\,\mathrm{e}^{-\iint_0^t \phi(s,z+v(s,z)\lambda)N(\mathrm{d}z\mathrm{d}s)}\,\mathcal{Z}_t^{v\lambda}\,\big] = \mathbb{E}\big[\,\mathrm{e}^{Y_t}\,\big]$$

where with $\phi^\lambda(t,z) = \phi(t,z+v(t,z)\lambda)$ and dropping the superscripts of $J^{v\lambda}$

$$Y_t = \iint_0^t \big(\log(J) - \phi^\lambda\big)(s,z)\ \tilde{N}(\mathrm{d}z\mathrm{d}s) - \iint_0^t \big(J + \phi^\lambda - 1 - \log(J)\big)(s,z)\nu(\mathrm{d}z)\mathrm{d}s\ .$$

Then Itô's formula gives

$$\begin{aligned}
\mathrm{d}\left[\,\mathrm{e}^Y\right]_t &= \mathrm{e}^Y \int \big[\mathrm{e}^{\log J - \phi^\lambda} - 1\big](t,z)\tilde{N}(\mathrm{d}z\mathrm{d}t) \\
&\quad + \mathrm{d}t\,\mathrm{e}^Y \int \big[\mathrm{e}^{\log J - \phi^\lambda} - 1 - (\log J - \phi^\lambda)\big](t,z)\ \nu(\mathrm{d}z) \\
&\quad - \mathrm{d}t\,\mathrm{e}^Y \int \big[J + \phi^\lambda - 1 - \log J\big](t,z)\ \nu(\mathrm{d}z) \\
&= \mathrm{e}^Y \int \big[J\,\mathrm{e}^{-\phi^\lambda} - 1\big](t,z)\tilde{N}(\mathrm{d}z\mathrm{d}t) \\
&\quad + \mathrm{d}t\,\mathrm{e}^Y \int \big[J\,\mathrm{e}^{-\phi^\lambda} - J\big](t,z)\ \nu(\mathrm{d}z)\ .
\end{aligned}$$

Since $\phi \in L^1(\mathrm{d}\nu \times \mathrm{d}t)$ the local martingale part is a true martingale. Taking expectations and substituting J defined in (3.6) and $\nu(\mathrm{d}z) = \varrho(z)\mathrm{d}z$ we are left with

$$\begin{aligned}
\mathbb{E}\big[\,\mathrm{e}^{Y_t}\,\big] &= 1 + \mathbb{E}\Big[\int_0^t \mathrm{e}^{Y_s} \int \Big[(\mathrm{e}^{-\phi^\lambda} - 1)\det(\mathrm{Id} + \nabla_z v\lambda)\frac{\varrho^{v\lambda}}{\varrho}\Big](s,z)\ \nu(\mathrm{d}z)\mathrm{d}s\,\Big] \\
&= 1 + \mathbb{E}\Big[\int_0^t \mathrm{e}^{Y_s} \int \big[(\mathrm{e}^{-\phi^\lambda} - 1)\varrho^{v\lambda}\big]\det(\mathrm{Id} + \nabla_z v\lambda)(s,z)\ \mathrm{d}z\mathrm{d}s\,\Big] \\
&= 1 + \int_0^t \mathbb{E}\big[\,\mathrm{e}^{Y_s}\,\big] \int \big(\mathrm{e}^{-\phi(s,z)} - 1\big)\ \varrho(z)\mathrm{d}z\mathrm{d}s\ .
\end{aligned}$$

The last equality follows substituting $z \mapsto z + v(\omega,t,z)\lambda$ for every ω, t, z where the corresponding Jacobian is given by (3.5). Solving the ordinary differential equation for $\mathbb{E}\big[\,\mathrm{e}^{Y_t}\,\big]$ we deduce

$$\mathbb{E}^\lambda\big[\mathrm{e}^{-\iint_0^t \phi(s,z)[N\circ\mathcal{T}_\lambda^\theta](\mathrm{d}z\mathrm{d}s)}\big] = \mathbb{E}\big[\mathrm{e}^{-\iint_0^t \phi(s,z)\ N(\mathrm{d}z\mathrm{d}s)}\big] = \exp\Big\{\iint_0^t \big(\mathrm{e}^{-\phi(s,z)} - 1\big)\ \nu(\mathrm{d}z)\mathrm{d}s\Big\}\ .$$

■

The following proposition shows that the family of Girsanov densities $(\mathcal{Z}^\lambda)_{\lambda\in\Lambda}$ in fact possesses an L^p derivative at $\lambda = 0$. Because of its significance as an adjoint operator in the integration-by-parts settings of the next chapter we will give it a special symbol. We denote for every $t \in [0,T]$

$$\delta_t(\theta) := \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\mathcal{Z}_t^\lambda \in \mathbb{R}^d\ . \tag{3.16}$$

Proposition 3.6. *The family of martingales $(\mathcal{Z}^\lambda)_{\lambda\in\Lambda}$ has an L^p-derivative at $\lambda = 0$ for every $1 \leqslant p \leqslant \infty$. It is given by*

$$\boldsymbol{\delta}_t(\theta) := \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\mathcal{Z}_t^\lambda = -\int_0^t h_s^* \mathrm{d}W_s + \iint_0^t \frac{\mathbf{div}_z\, \varrho v}{\varrho}\tilde{N}(\mathrm{d}z\mathrm{d}s) \ . \tag{3.17}$$

Proof. For simplicity we only prove the case of $p = 2$. This means we show that for any $t \geqslant 0$

$$\mathbb{E}\big[\ |\mathcal{Z}_\cdot^\lambda - 1 - \boldsymbol{\delta}_\cdot(\theta)\lambda|^2_{[0,t]}\ \big] = o(|\lambda|) \ , \tag{3.18}$$

where $\boldsymbol{\delta}_t(\theta)$ is defined as the right-hand-side of (3.17). Recall that for any t $\mathcal{Z}_t^\lambda = \mathcal{Z}_t^{h\lambda} \times \mathcal{Z}_t^{v\lambda}$ where the components are given by (3.13) (resp. (3.14)), each satisfying (3.12) with $v \equiv 0$ (resp. $h \equiv 0$). By the product rule we have

$$\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\mathcal{Z}_t^\lambda = \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\mathcal{Z}_t^{h\lambda} + \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\mathcal{Z}_t^{v\lambda} \ , \tag{3.19}$$

such that we can investigate the two components individually. Let us investigate the derivative on Wiener space. Using (3.13) and Itô's isometry twice we obtain

$$\begin{aligned}
\mathbb{E}\big[\ |\mathcal{Z}_t^{h\lambda} - 1 + \int_0^t (h_s\lambda)^*\mathrm{d}W_s|^2\ \big] &= \mathbb{E}\big[\ |\int_0^t (\mathcal{Z}_s^{h\lambda} - 1)(h_s\lambda)^*\mathrm{d}W_s|^2\ \big] \\
&= \mathbb{E}\big[\ |\int_0^t \big(\int_0^s \mathcal{Z}_r^{h\lambda}(h_r\lambda)^*\mathrm{d}W_r\big)(h_s\lambda)^*\mathrm{d}W_s|^2\ \big] \\
&= \mathbb{E}\big[\ \int_0^t \int_0^s \big|\mathcal{Z}_r^{h\lambda}(h_r\lambda)^*\mathrm{d}W_r\big|^2 \big(\sum_{j=1}^m (h_s^{i\cdot})^*\lambda\big)^2\ \mathrm{d}s\ \big] \\
&= \int_0^t \mathbb{E}\big[\ \int_0^s (\mathcal{Z}_r^{h\lambda})^2 \big(\sum_{j=1}^m (h_s^{i\cdot})^*\lambda\big)^2 \big(\sum_{j=1}^m (h_r^{i\cdot})^*\lambda\big)^2\ \mathrm{d}r\ \big]\mathrm{d}s \\
&\lesssim |\lambda|^4 \mathbb{E}\big[\ \int_0^t \int_0^s (\mathcal{Z}_r^{h\lambda})^2\ \mathrm{d}r\mathrm{d}s|\ \big] \leqslant |\lambda|^4 \frac{t^2}{2}\mathbb{E}\big[\ |\mathcal{Z}_\cdot^{h\lambda}|^2_{[0,t]}\ \big] \ ,
\end{aligned}$$

where the first inequality is due to the boundedness of $h \in \mathcal{H}_0$. Clearly this estimate holds uniformly over the interval $[0, t]$.
Now we turn to $\mathcal{Z}^{v\lambda}$. Using (3.14) and Itô's isometry we obtain

$$\begin{aligned}
&\mathbb{E}\big[\ |\mathcal{Z}_t^{v\lambda} - 1 - \iint_0^t \varrho^{-1}\,\mathbf{div}_z(v\varrho)\lambda\ \tilde{N}(\mathrm{d}z\mathrm{d}s)|^2\ \big] \\
&\quad = \mathbb{E}\big[\ \iint_0^t \mathcal{Z}_{s-}^{v\lambda}\big(J^{v\lambda}(s,z) - 1\big) - \varrho^{-1}\,\mathbf{div}_z(v\varrho)\lambda\ \tilde{N}(\mathrm{d}z\mathrm{d}s)|^2\ \big] \\
&\quad = \mathbb{E}\big[\ \iint_0^t \big(\mathcal{Z}_{s-}^{v\lambda}\big(J^{v\lambda}(s,z) - 1\big) - \varrho^{-1}\,\mathbf{div}_z(v\varrho)\lambda\big)^2\ \nu(\mathrm{d}z)\mathrm{d}s\ \big] \\
&\quad \leqslant 2\mathbb{E}\big[\ \iint_0^t \big(\mathcal{Z}_{s-}^{v\lambda} - 1\big)^2\big(J^{v\lambda}(s,z) - 1\big)^2\ \nu(\mathrm{d}z)\mathrm{d}s\ \big] \\
&\qquad + 2\mathbb{E}\big[\ \iint_0^t \big(J^{v\lambda}(s,z) - 1 - \varrho^{-1}\,\mathbf{div}_z(v\varrho)\lambda\big)^2\ \nu(\mathrm{d}z)\mathrm{d}s\ \big] \ .
\end{aligned}$$

We estimate the first of the two expectations further, again relying on (3.14).

$$\begin{aligned}
&\mathbb{E}\Big[\iint_0^t \left(\mathcal{Z}_{s-}^{v\lambda} - 1\right)^2 \left(J^{v\lambda}(s,z) - 1\right)^2 \ \nu(\mathrm{d}z)\mathrm{d}s \ \Big] \\
&= \mathbb{E}\Big[\iint_0^t \left(\iint_0^s \mathcal{Z}_{r-}^{v\lambda} \left(J^{v\lambda}(r,y) - 1\right) \ \tilde{N}(\mathrm{d}y\mathrm{d}r) \right)^2 \left(J^{v\lambda}(s,z) - 1\right)^2 \ \nu(\mathrm{d}z)\mathrm{d}s \ \Big] \\
&= \iint_0^t \mathbb{E}\Big[\iint_0^s \left(\mathcal{Z}_{r-}^{v\lambda}\right)^2 \left(J^{v\lambda}(r,y) - 1\right)^2 \left(J^{v\lambda}(s,z) - 1\right)^2 \nu(\mathrm{d}y)\mathrm{d}r \ \Big] \ \nu(\mathrm{d}z)\mathrm{d}s \ .
\end{aligned}$$

This time changing the order of integration with respect to $\mathbb{P}$ and $\nu \otimes \mathrm{d}s$ to apply Itô's isometry a second time is slightly more delicate then it was above since the Lévy measure is not a finite measure. However Corollary 3.3 guarantees sufficient integrability. In particular it further implies that the above expression is bounded by

$$\begin{aligned}
&|\lambda|^4 \iint_0^t |V'(z)|^2 \iint_0^s |V'(y)|^2 \mathbb{E}\Big[\ \left(\mathcal{Z}_{r-}^{v\lambda}\right)^2 \ \Big] \nu(\mathrm{d}y)\mathrm{d}r \ \nu(\mathrm{d}z)\mathrm{d}s \\
&\qquad \lesssim |\lambda|^4 \frac{t^2}{2} \mathbb{E}\Big[\ |\mathcal{Z}_{\cdot}^{v\lambda}|_{[0,t]}^2 \ \Big] \ .
\end{aligned}$$

It remains to consider the second of the two expectations. Under the expectation we have

$$\iint_0^t |J^{v\lambda}(s,z) - 1 - \varrho^{-1} \, \mathbf{div}_z(\varrho v)\lambda|^2 \nu(\mathrm{d}z)\mathrm{d}s \tag{3.20}$$

$$= \iint_0^t | \int_0^1 \left\{ \tfrac{\partial}{\partial \lambda_j} J^{v\lambda}\big|_{\lambda = \tau\lambda} \lambda - (\mathbf{div}_z \, v)\lambda - \frac{\nabla \varrho}{\varrho} v\lambda \right\} \ \mathrm{d}\tau|^2 \nu(\mathrm{d}z)\mathrm{d}s \ . \tag{3.21}$$

Recall (3.8)

$$\tfrac{\partial}{\partial \lambda_j} J^{v\lambda} = \det(\mathrm{Id} + \nabla v\lambda) \Big(\mathrm{Tr}\big((\mathrm{Id} + \nabla v\lambda)^{-1} \nabla v e_j \big) + \frac{\nabla \varrho(\cdot + v\lambda) v e_j}{\varrho} \Big) \tag{3.22}$$

with $e_1, \ldots, e_m$ the canonical basis of $\mathbb{R}^m$. By definition $(\mathbf{div}_z \, v)e_j = \mathrm{Tr} \, \nabla v e_j$ such that

$$\begin{aligned}
|\lambda_j \, \mathrm{Tr}\left((\mathrm{Id} + \nabla v\lambda)^{-1} \nabla v e_j \right) - \lambda_j (\mathbf{div}_z \, v) e_j| &= |\lambda_j \, \mathrm{Tr}\left(\big[(\mathrm{Id} + \nabla v\lambda)^{-1} - \mathrm{Id} \big] \nabla v e_j \right)| \\
&\leqslant \|\nabla v\|^2 |\lambda| |\lambda_j|,
\end{aligned}$$

which we sum up over j to obtain with the boundedness of $v \in \mathcal{V}_0$ a bound of order $V^2|\lambda|^2$. Also

$$|\frac{\nabla \varrho(\cdot + v\lambda)}{\varrho} v\lambda - \frac{\nabla \varrho}{\varrho} v\lambda| \lesssim \frac{|v\lambda|^\delta}{\varrho} |v\lambda| \leqslant \frac{V^{1+\delta}}{\varrho} |\lambda|^{1+\delta} \ .$$

Hence we have the estimate

$$2\mathbb{E}\Big[\iint_0^t \left(J^{v\lambda}(s,z) - 1 - \varrho^{-1} \, \mathbf{div}_z(v\varrho)\lambda \right)^2 \nu(\mathrm{d}z)\mathrm{d}s \ \Big] \lesssim |\lambda|^{2+2\delta}$$

Combining these 3 estimates we obtain that the left-hand-side of 3.18 is of order $O(|\lambda|^{2+2\delta})$. The proposition is proven. ∎

3.1.3 The integration-by-parts formula

We have established the ray-continuity of the measure $\mathbb{P}$ in the direction of shifts $\theta \in \Theta_0$ and the differentiability of the Girsanov density $\mathcal{Z}^{\theta,\lambda}$.
The integration-by-parts formula then follows along the arguments of Section 3.1.1. Indeed, if a functional $\Psi \in L^p(\Omega,\mathbb{P})$ for some $p > 1$ has an L^p-derivative in the direction of $\theta \in \Theta_0$ we deduce by Proposition 3.5 together with the product rule (2.5) that

$$0 = \mathbb{E}\big[\ \tfrac{\partial}{\partial\lambda}\big|_{\lambda=0} \left\{\Psi \circ \mathcal{T}^\theta_\lambda\ \mathcal{Z}^\lambda_t\right\}\ \big] = \mathbb{E}\big[\ \Psi(\tfrac{\partial}{\partial\lambda}\big|_{\lambda=0}\mathcal{Z}^\lambda_t) + (\mathscr{D}_\theta\Psi)\mathcal{Z}^0_t\ \big]\ , \tag{3.23}$$

which then gives the *integration-by-parts* formula

$$\mathbb{E}\big[\mathscr{D}_\theta\Psi\big] = -\mathbb{E}\big[\Psi\boldsymbol{\delta}_t(\theta)\big] = \mathbb{E}\Big[\Psi\left\{\int_0^t h^*_s \mathrm{d}W_s - \iint_0^t \varrho^{-1}\,\mathbf{div}_z(v\varrho)(s,z)\ \tilde{N}(\mathrm{d}z\mathrm{d}s)\right\}\Big]\ . \tag{3.24}$$

We may now extend the integration-by-parts formula to a larger set of perturbations Θ as long as the right hand side is well defined. Indeed we may consider the following.

Condition 8 (Admissible perturbation)**.** We call a perturbation θ ADMISSIBLE if it is a pair $\theta = (h,v) \in \Theta_\varrho := \mathcal{H} \times \mathcal{V}_\varrho$ where

- $\mathcal{H}$ is the closure of $\mathcal{H}_0$ with respect to the $L^2(\mathbb{P}\times \mathrm{d}t)$ norm

$$\|h\|_{\mathcal{H}} := \Big(\int_0^t \mathbb{E}\big[\ \|h_s(\omega)\|^2\ \big]\mathrm{d}s\Big)^{\frac{1}{2}}\ . \tag{3.25}$$

- $\mathcal{V}_\varrho$ is the set closure of $\mathcal{V}_0$ with respect to the weighted norm[1]

$$\|v\|_{\mathcal{V}_\varrho} := \Big(\iint_0^t \mathbb{E}\big[|\,\mathbf{div}_z\, v(\omega,s,z)|^2 + |(\nabla\log\varrho(z))v(\omega,s,z)|^2\big]\nu(\mathrm{d}z)\mathrm{d}s\Big)^{\frac{1}{2}}\ . \tag{3.26}$$

Remark 3.7. A similar norm to $\|\cdot\|_{\mathcal{V}_\varrho}$ appears in [Son14] for deterministic perturbations where additive Poisson noise without a Gaussian component is considered.

It is clear that if (3.25) and (3.26) are finite the right hand side of (3.24) is also finite. Moreover if we consider $\boldsymbol{\delta}_t$ given by (3.17) as a linear operator it is closable in $L^2(\Omega)$.

Lemma 3.8. *There exists a closed linear extension of the linear operator*

$$\boldsymbol{\delta}_t : \Theta_0 \to L^2(\Omega) \tag{3.27}$$

to Θ_ϱ, i.e. $\boldsymbol{\delta}_t$ is closable and its extension is again denoted by $\boldsymbol{\delta}_t$.

[1]Here we tacitly assume that the Lévy density ϱ is not piecewise constant.

Proof. It is sufficient to show that if for a sequence $(\theta_n)_{n\in\mathbb{N}} \subset \Theta_0$ we have $\theta_n \to 0$ in Θ_ϱ then $\lim_{n\to\infty} \boldsymbol{\delta}_t(\theta_n) = 0$ in $L^2(\Omega)$ (*cf.* [Yos80, II.6, Prop.2, p.77]). Indeed, the sequence $\boldsymbol{\delta}(\theta_n)$ is a Cauchy sequence. Denote $\theta_n = (h_n, v_n), n \in \mathbb{N}$. Then for $n, m \in \mathbb{N}$ we have

$$\begin{aligned}
&\mathbb{E}\big[\ |\boldsymbol{\delta}(\theta_m) - \boldsymbol{\delta}(\theta_n)|^2\ \big] \\
&\qquad = \mathbb{E}\Big[\ \Big| - \int_0^t [h_n - h_m](s)^* \mathrm{d}W_s + \iint_0^t \varrho^{-1}\, \mathbf{div}_z([v_n - v_m]\varrho)(s,z)\ \tilde{N}(\mathrm{d}z\mathrm{d}s)\Big|^2\ \Big] \\
&\qquad \leqslant \|h_n - h_m\|^2_{\mathcal{H}} + \|v_n - v_m\|^2_{\mathcal{V}_\varrho}\ .
\end{aligned}$$

■

The next lemma demonstrates that we can take limits in (3.24).

Lemma 3.9. *Consider $\theta = (h, v)$ in Θ_ϱ and an approximating sequence of simple perturbations $\theta_n = (h_n, v_n) \in \Theta_0$. Assume that a random variable $\Psi \in L^2(\Omega, \mathbb{P})$ is L^2-differentiable with respect to Θ_0. Then the limit $\lim_{n\to\infty} \mathbb{E}\big[\ \mathscr{D}_{\theta_n}\Psi\ \big]$ exists and is independent of the choice of the approximating sequence.*

Proof. Existence follows by the linearity of the operations involved and the Cauchy–Schwartz inequality. Indeed we have

$$\begin{aligned}
\mathbb{E}\big[\ \mathscr{D}_{\theta_n}\Psi - \mathscr{D}_{\theta_m}\Psi\ \big] &= -\mathbb{E}\big[\ \Psi\boldsymbol{\delta}(\theta_n - \theta_m)\ \big] \\
&\leqslant \mathbb{E}\big[\ \Psi^2\ \big]^{\frac{1}{2}} \Big\{\|h_n - h_m\|^2_{\mathcal{H}} + \|v_n - v_m\|^2_{\mathcal{V}_\varrho}\Big\}^{\frac{1}{2}}\ .
\end{aligned}$$

We repeat the estimation in reverse order to see that $\mathbb{E}\big[\ \mathscr{D}_{\theta_m}\Psi - \mathscr{D}_{\theta_n}\Psi\ \big] \to 0$ as m, n go to infinity. ■

So far we know that we can take the limit of (3.24) for any approximating sequence of $\theta \in \Theta_\varrho$. It is however not clear whether and in which sense the limit $\lim_{n\to\infty} \mathscr{D}_{\theta_n}\Psi$ exists. We therefore concentrate on $\theta \in \Theta_\varrho \cap \Theta_p$ where we are able to guarantee the existence of an L^p-derivative by Theorem 2.13 for random variables Ψ of interest.
We reformulate the integration-by-parts formula in a "ready to use" way for random weights Ψ, i.e. for products of the form $\Psi f(\xi)$ – quite in the spirit of the "integration-by-parts-setting" of [BGJ87] (*cf.* also [Mal97, Prop. 1.3.3, p.68]).

Theorem 3.10 (integration-by-parts-setting)**.** *Let ξ be the solution to the SDE 2.1 and let $\Psi \in L^2(\Omega, \mathbb{P})$ be L^2-differentiable in the direction of $\theta \in \Theta_\varrho \cap \Theta_2$. Assume further that there exists a sequence $\theta_n \in \Theta_0$ such that $\theta_n \to \theta$ in Θ_ϱ and that for $t \in [0, T]$ we have*

$$\mathscr{D}_{\theta_n}\xi_t \to \mathscr{D}_\theta\xi_t \text{ in } L^2\ . \tag{3.28}$$

Then we have for any $f \in \mathscr{C}^1_b(\mathbb{R}^d)$

$$\begin{aligned}
\mathbb{E}\big[\ \Psi\nabla f(\xi_t)\mathscr{D}_\theta\xi_t\ \big] &= -\mathbb{E}\Big[\ f(\xi_t)\Big\{\Psi\boldsymbol{\delta}(\theta) + \mathscr{D}_\theta\Psi\Big\}\ \Big] \\
&= -\mathbb{E}\Big[\ f(\xi_t)\Big\{\Psi\Big(-\int_0^t h^*_s \mathrm{d}W_s + \iint_0^t \frac{\mathbf{div}_z\,\varrho v}{\varrho}\ \tilde{N}(\mathrm{d}z\mathrm{d}s)\Big) + \mathscr{D}_\theta\Psi\Big\}\ \Big] \qquad (3.29) \\
&= -\mathbb{E}\big[\ f(\xi_t)\Gamma_\theta(\Psi)\ \big]\ .
\end{aligned}$$

Proof. For any of the approximating $\theta_n \in \Theta_0$ the equation is just an application of the chain rule (Proposition 2.3). Indeed, the random variable $\Phi := \Psi f(\xi_t)$ is L^2-differentiable with

$$\mathscr{D}_{\theta_n}\Phi = \Psi \nabla f(\xi_t)\mathscr{D}_{\theta_n}\xi_t + f(\xi_t)\mathscr{D}_{\theta_n}\Psi \ .$$

Taking expectations and making use of (3.24) we obtain

$$-\mathbb{E}\big[\ \Psi f(\xi_t)\boldsymbol{\delta}(\theta_n)\ \big] = \mathbb{E}\big[\ \Psi\nabla f(\xi_t)\mathscr{D}_{\theta_n}\xi_t\ \big] + \mathbb{E}\big[\ f(\xi_t)\mathscr{D}_{\theta_n}\Psi\ \big] \ . \tag{3.30}$$

Hence we obtained (3.29). Since $\boldsymbol{\delta}$ is closable in L^2 by Lemma 3.8 it follows by the Cauchy–Schwartz-inequality with the boundedness of f that

$$\mathbb{E}\big[\ f(\xi_t)\Psi\boldsymbol{\delta}(\theta_n)\ \big] \to \mathbb{E}\big[\ f(\xi_t)\Psi\boldsymbol{\delta}(\theta)\ \big] \ .$$

Furthermore the L^2-convergence of $\mathscr{D}_{\theta_n}\xi_t$ and a similar Cauchy–Schwartz argument together with the boundedness of ∇f imply that also

$$\mathbb{E}\big[\ \nabla f(\xi_t)\Psi\mathscr{D}_{\theta_n}\xi_t\ \big] \to \mathbb{E}\big[\ \nabla f(\xi_t)\Psi\mathscr{D}_{\theta}\xi_t\ \big] \ .$$

Now since two out of the three terms in (3.30) converge and therefore also

$$\mathbb{E}\big[\ f(\xi_t)\mathscr{D}_{\theta_n}\Psi\ \big] \to \mathbb{E}\big[\ f(\xi_t)\mathscr{D}_{\theta}\Psi\ \big] \ .$$

Hence we have proven (3.29). ■

Remark 3.11. If $\theta \in \Theta_2$ is only a perturbation of the Wiener measure ($v \equiv 0$) then the convergence of $\mathscr{D}_{\theta_n}\xi_t$ to $\mathscr{D}_{\theta}\xi_t$ in L^2 follows from the fact that $h_n \to h$ in $\mathcal{H}$. Indeed, by Theorem 2.13 the dynamic of the difference is governed by

$$\begin{aligned}\mathrm{d}[\mathscr{D}_\theta\xi - \mathscr{D}_{\theta_n}\xi]_t = {}& \nabla\beta(\xi_{t-})[\mathscr{D}_\theta\xi - \mathscr{D}_{\theta_n}\xi]_t\mathrm{d}t + \nabla\sigma(\xi_{t-})[\mathscr{D}_\theta\xi - \mathscr{D}_{\theta_n}\xi]_t\mathrm{d}W_t \\ &+ \sigma(\xi_{t-})[h-h_n](t)\mathrm{d}t \ .\end{aligned}$$

Since all coefficients and their derivatives are bounded we can estimate the mean squared deviation with the help of Itô's isometry to obtain

$$\mathbb{E}\big[\ \|\mathscr{D}_\theta\xi - \mathscr{D}_{\theta_n}\xi\|^2_{[0,T]}\ \big] \lesssim \int_0^T \mathbb{E}\big[\|\mathscr{D}_\theta\xi - \mathscr{D}_{\theta_n}\xi\|^2_{[0,t]}\mathrm{d}t\big] + \int_0^T \mathbb{E}\big[\|[h-h_n](t)\|^2\big]\mathrm{d}t \ .$$

Remark 3.12. If we consider the identity function $f(x) = x$ and the constant functional $\Psi \equiv \mathbf{1}$, the integration-by-parts formula reads

$$\begin{aligned}\mathbb{E}\big[\ \mathscr{D}_\theta\xi_t\ \big] &= -\mathbb{E}\Big[\ \xi_t\Big\{-\int_0^t h_s^*\mathrm{d}W_s + \iint_0^t \frac{\mathbf{div}_z\, \varrho v}{\varrho}\ \tilde{N}(\mathrm{d}z\mathrm{d}s)\Big\}\ \Big] \\ &= -\mathbb{E}\big[\ \xi_t\Gamma_\theta(\mathbf{1})\ \big] \ .\end{aligned} \tag{3.31}$$

This implies that the random variable $-\Gamma_\theta(\mathbf{1})$ is the Radon–Nikodým density of the measure $\partial_\theta\mu$ with respect to the measure μ, the law of $\xi_t(x)$ on $\mathbb{D}[0,T]$, the so called *logarithmic derivative* (*cf.* [Bog10]).

For convenience we formulate two corollaries for vector- and matrix weights. In analogy to the standard divergence we introduce the operator

$$\mathrm{div}_\theta \boldsymbol{\Psi} := \mathrm{Tr}\, \mathscr{D}_\theta \boldsymbol{\Psi} \ , \tag{3.32}$$

for differentiable $\boldsymbol{\Psi} \in L^p(\Omega, \mathbb{P}; \mathbb{R}^d), p \geqslant 1$.

Corollary 3.13 (vector version)**.** *Let $\boldsymbol{\Psi} \in L^p(\Omega, \mathbb{P}; \mathbb{R}^d), \forall p \geqslant 1$ be a random vector in $\mathbb{R}^d$ which is L^p-differentiable in the direction of $\theta \in \Theta_\varrho \cap \Theta_p, \forall p \geqslant 1$ and $f \in \mathscr{C}_b^1(\mathbb{R}^d)$. Furthermore assume that we have a sequence $(\theta_n)_{n\in\mathbb{N}} \subset \Theta_0$ satisfying* (3.28) *for all coordinates of Ψ. Then we have*

$$\begin{aligned}\mathbb{E}\big[\ \nabla f(\xi_t)\mathscr{D}_\theta\xi_t\boldsymbol{\Psi}\ \big] = -\mathbb{E}\Big[f(\xi_t)\Big\{\Big\langle \boldsymbol{\Psi}, \Big(-\int_0^t h_s^* \mathrm{d}W_s + \iint_0^t \frac{\mathbf{div}_z\, \varrho v}{\varrho}\ \tilde{N}(\mathrm{d}z\mathrm{d}s)\Big)\Big\rangle \\ + \mathrm{div}_\theta\, \boldsymbol{\Psi}\Big\}\Big]\ .\end{aligned} \tag{3.33}$$

We also write

$$\mathbb{E}\big[\ \nabla f(\xi_t)\mathscr{D}_\theta\xi_t\boldsymbol{\Psi}\ \big] = -\mathbb{E}\big[\ f(\xi_t)\boldsymbol{\Gamma}_\theta\left(\boldsymbol{\Psi}\right)\ \big]\ , \tag{3.34}$$

with

$$\begin{aligned}\boldsymbol{\Gamma}_\theta\left(\boldsymbol{\Psi}\right) &= \Big\langle \boldsymbol{\Psi}, \boldsymbol{\delta}(\theta)\Big\rangle + \mathrm{div}_\theta\, \boldsymbol{\Psi} \\ &= \Big\langle \boldsymbol{\Psi}, \Big(-\int_0^t h_s^* \mathrm{d}W_s + \iint_0^t \frac{\mathbf{div}_z\, \varrho v}{\varrho}\ \tilde{N}(\mathrm{d}z\mathrm{d}s)\Big)\Big\rangle + \mathrm{Tr}\Big(\mathscr{D}_\theta\boldsymbol{\Psi}\Big)\ .\end{aligned} \tag{3.35}$$

Proof. It is sufficient to notice that

$$\begin{aligned}\mathbb{E}\big[\ \nabla f(\xi_t)\mathscr{D}_\theta\xi_t\boldsymbol{\Psi}\ \big] &= \sum_{i=1}^d \mathbb{E}\big[\ \boldsymbol{\Psi}^i\Big(\nabla f(\xi_t)\mathscr{D}_\theta\xi_t\Big)^i\ \big] = -\sum_{i=1}^d \mathbb{E}\big[\ f(\xi_t)\Gamma_\theta^i(\boldsymbol{\Psi}^i)\ \big] \\ &= -\mathbb{E}\big[\ f(\xi_t)\Big\{\sum_{i=1}^d \boldsymbol{\Psi}^i\Big(-\int_0^t h_s^*\mathrm{d}W_s + \iint_0^t \frac{\mathbf{div}_z\, \varrho v}{\varrho}\ \tilde{N}(\mathrm{d}z\mathrm{d}s)\Big)^i + \Big(\mathscr{D}_\theta\boldsymbol{\Psi}^i\Big)^i\Big\}\ \big]\ .\end{aligned}$$

■

Analogously we also obtain a matrix formulation. For differentiable $\boldsymbol{\Psi} \in L^p(\Omega, \mathbb{P}; \mathbb{R}^{d\times d})$, $p \geqslant 1$ we write

$$\mathbf{div}_\theta\, \boldsymbol{\Psi} := \mathbf{Tr}\,|_{\mathbb{R}^d}\mathscr{D}_\theta\boldsymbol{\Psi} \tag{3.36}$$

where $\mathbf{Tr}\,|_{\mathbb{R}^d} : \mathbb{R}^d \otimes \mathbb{R}^d \otimes \mathbb{R}^d \mapsto \mathbb{R}^{d*}$ is the tensor contraction (trace)[2] to $\mathbb{R}^{d*}$ and $\mathbf{Tr}\,|_{\mathbb{R}^d}\mathscr{D}_\theta\boldsymbol{\Psi}$ is the row vector $(\mathrm{Tr}\,\mathscr{D}_\theta\boldsymbol{\Psi}_{\cdot j})^{(j)}$ obtained by column wise application.

[2] see [MRA07, p.344] or [BG80, pp.85-86]

Corollary 3.14 (matrix version)**.** *Let $\mathbf{\Psi} \in L^p(\Omega, \mathbb{P}; \mathbb{R}^{d\times d}), \forall p \geqslant 1$ be a random vector in $\mathbb{R}^d$ which is L^p-differentiable in the direction of $\theta \in \Theta_\varrho \cap \Theta_p, \forall p \geqslant 1$ and $f \in \mathscr{C}_b^1(\mathbb{R}^d)$. Furthermore assume that we have a sequence $(\theta_n)_{n\in\mathbb{N}} \subset \Theta_0$ satisfying* (3.28) *for all coordinates of Ψ. Then we have*

$$\begin{aligned}\mathbb{E}\big[\ \nabla f(\xi_t)\mathscr{D}_\theta\xi_t\mathbf{\Psi}\ \big] = -\mathbb{E}\Big[f(\xi_t)\Big\{\mathbf{\Psi}^*\Big(-\int_0^t h_s^*\mathrm{d}W_s + \iint_0^t \frac{\mathbf{div}_z\,\varrho v}{\varrho}\ \tilde{N}(\mathrm{d}z\mathrm{d}s)\Big) \\ + \mathbf{div}_\theta\,\mathbf{\Psi}\Big\}\Big]\ .\end{aligned} \tag{3.37}$$

We also write

$$\mathbb{E}\big[\ \nabla f(\xi_t)\mathscr{D}_\theta\xi_t\mathbf{\Psi}\ \big] = -\mathbb{E}\big[\ f(\xi_t)\mathbf{\Gamma}_\theta\left(\mathbf{\Psi}\right)\ \big]\ , \tag{3.38}$$

with

$$\begin{aligned}\mathbf{\Gamma}_\theta\left(\mathbf{\Psi}\right) &= \mathbf{\Psi}^*\delta(\theta) + \mathbf{div}_\theta\,\mathbf{\Psi} \\ &= \mathbf{\Psi}^*\Big(-\int_0^t h_s^*\mathrm{d}W_s + \iint_0^t \frac{\mathbf{div}_z\,\varrho v}{\varrho}\ \tilde{N}(\mathrm{d}z\mathrm{d}s)\Big) + \mathbf{div}_\theta\,\mathbf{\Psi}\ .\end{aligned} \tag{3.39}$$

3.2 Application of the integration-by-parts formula

We are now in the situation to apply the integration-by-parts formula. Recall that by Theorem 2.14 we have the equality $\mathscr{D}_\theta\xi_t = \nabla\xi_t\mathscr{A}_t^\theta$.
The trick is now to find a "good" perturbation that makes the reduced Malliavin matrix $\mathscr{A}_t^\theta$ invertible.

3.2.1 Gradient estimates and the Strong Feller property

Let ξ be the unique strong solution to (2.1). As a direct application of the integration-by-parts formula we obtain an estimate of the gradient of the semigroup $\mathcal{P} = (\mathcal{P}_t)_{t\geqslant 0}$ generated by the Markov process ξ by the relation

$$\mathcal{P}_t f(x) = \mathbb{E}\big[\ f\big(\xi_t(x)\big)\ \big]\ , \quad t \geqslant 0, x \in \mathbb{R}^d, f \in \mathscr{B}_b(\mathbb{R}^d)\ . \tag{3.40}$$

The domain of $\mathcal{P}$ is the set of real valued bounded measurable functions $\mathscr{B}_b(\mathbb{R}^d)$. We are interested in the following smoothing property of $\mathcal{P}$.

Definition 3.15. The Markov semigroup $\mathcal{P} = (\mathcal{P}_t)_{t\geqslant 0}$ is said to have the STRONG FELLER property if it maps $\mathscr{B}_b(\mathbb{R}^d)$ to $\mathscr{C}_b(\mathbb{R}^d)$.

Observe that if $f \in \mathscr{C}_b^1(\mathbb{R}^d)$ is smooth and bounded its image under the semigroup is differentiable. Indeed, with Lebesgue's dominated convergence theorem we deduce

$$\nabla\mathcal{P}_t f(x) = \mathbb{E}\big[\ \nabla f(\xi_t)\nabla\xi_t\ \big]\ .$$

In view of Theorem 2.14 and Corollary 3.14 we would like to choose a matrix $\mathbf{\Psi} = (\mathscr{A}_t^\theta)^{-1}$ to obtain a uniform estimate. Indeed we have the following.

Lemma 3.16. *Let $\theta \in \Theta_\varrho \cap \Theta_2$. Assume that $\mathbf{\Psi} = (\mathscr{A}_t^\theta)^{-1}$ exists and is in $L^2(\Omega)$ and that it is L^2-differentiable in the direction of θ. Assume also that there exists a simple approximation $(\theta_n)_{n\in\mathbb{N}} \subset \Theta_0$ satisfying* (3.28)*. Then*

$$\nabla \mathcal{P}_t f(x) = -\mathbb{E}\big[\ f(\xi_t)\mathbf{\Gamma}_\theta\left((\mathscr{A}_t^\theta)^{-1}\right)\ \big]\ , \quad f \in \mathscr{C}_b^\infty(\mathbb{R}^d)\ , \tag{3.41}$$

where $\mathbf{\Gamma}_\theta$ is given by (3.39)*.*

Proof. A direct application of Corollary 3.14. ■

Theorem 3.17. *Under the conditions of Lemma 3.16 we have*

$$|\nabla \mathcal{P}_t f(x)| \lesssim \|f\|_\infty\ , \quad f \in \mathscr{B}_b(\mathbb{R}^d)\ . \tag{3.42}$$

In particularly the semigroup $\mathcal{P}$ has the strong Feller property and the process ξ is a strong Feller process.

Proof. By Lemma 3.16 we have for any $f \in \mathscr{C}_b^\infty(\mathbb{R}^d)$

$$\begin{aligned}|\nabla \mathcal{P}_t f(x)| =& |\mathbb{E}\big[\ \nabla f(\xi_t)\nabla\xi_t\ \big]| = |\mathbb{E}\big[\ \nabla f(\xi_t)\mathscr{D}_\theta\xi_t(\mathscr{A}_t^\theta)^{-1}\ \big]| \\ \leqslant& \|f\|_\infty \mathbb{E}\big[\ \|\mathbf{\Gamma}_\theta\left((\mathscr{A}_t^\theta)^{-1}\right)\|\ \big]\ .\end{aligned}$$

If $f \in \mathcal{B}_b(\mathbb{R}^d)$ we may approximate it uniformly by $f_n \in \mathscr{C}_b^\infty(\mathbb{R}^d)$ (e.g. by convolution of f with a mollifier). Since

$$|\nabla \mathcal{P}_t[f_n - f_m](x)| \lesssim \|f_n - f_m\|_\infty$$

the limit exists for any x and we have (3.42). Hence $\mathcal{P}_t f$ is differentiable and consequently we obtain the strong Feller property

$$\mathcal{P}_t : \mathscr{B}_b(\mathbb{R}^d) \to \mathscr{C}_b(\mathbb{R}^d)\ .$$

■

3.2.2 Smooth densities of the law

The proof of the existence of a smooth density of the law of ξ_t for fixed $t \in [0, T]$ relies on the following observation from harmonic analysis. For a proof we refer to [Str81b, Lemma 3.1, p.56].

Lemma 3.18. *Let μ be a finite Radon measure on $\mathbb{R}^d$ and assume that there is $N \geqslant d$ such that for all multi-indices α with $|\alpha| \leqslant N$ we have*

$$\left|\int_{\mathbb{R}^d} \partial^\alpha f(x)\mu(\mathrm{d}x)\right| \lesssim \|f\|_\infty\ , \quad f \in \mathscr{C}_b^\infty(\mathbb{R}^d)\ . \tag{3.43}$$

Then μ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$ with a density of class $\mathscr{C}_b^k(\mathbb{R}^d)$ with $k = N - d - 1$.

Applying the Lemma to the law of ξ_t on $\mathbb{R}^d$ we want to choose $\mathbf{\Psi} = (\mathscr{D}_\theta \xi_t)^{-1}$ in (3.37). Indeed if we find an approximation satisfying (3.28) we obtain from Corollary 3.13

$$\mathbb{E}\big[\ \nabla f(\xi_t)\ \big] = -\mathbb{E}\big[\ f(\xi_t)\mathbf{\Gamma}_\theta\left((\mathscr{D}_\theta \xi_t)^{-1}\right)\ \big]\ , \quad f \in \mathscr{C}_b^\infty(\mathbb{R}^d)\ . \tag{3.44}$$

This guarantees the existence of a continuous (i.e. $\mathscr{C}_b^0(\mathbb{R}^d)$) density with respect to the Lebesgue measure.

Lemma 3.19. *Assume that $\mathbf{\Psi} = (\mathscr{D}_\theta \xi_t)^{-1} \in L^2(\Omega)$ with $\theta \in \Theta_\varrho \cap \Theta_2$ is L^2-differentiable in the direction of θ and there exists a simple approximation satisfying* (3.28)*. Then*

$$\mathbb{E}\big[\ \nabla f(\xi_t)\ \big] = -\mathbb{E}\big[\ f(\xi_t)\mathbf{\Gamma}_\theta\left((\mathscr{D}_\theta \xi_t)^{-1}\right)\ \big]\ , \quad f \in \mathscr{C}_b^\infty(\mathbb{R}^d)\ , \tag{3.45}$$

where $\mathbf{\Gamma}_\theta$ is given by (3.39)*.*

Proof. A direct application of Lemma 3.14. ∎

Since we are in the regime of ξ being a diffeomorphism (the Lévy measure is supported in a sufficiently small ball, Section 2.2.2, in particular Corollary 2.8) we remark that by the relation $\mathscr{D}_\theta \xi_t = \nabla \xi_t \mathscr{A}_t^\theta$ the invertibility of the Malliavin matrix can be deduced from the invertibility of the reduced Malliavin matrix. Indeed

$$\det(\mathscr{D}_\theta \xi_t)^{-1} = \det(\nabla \xi_t)^{-1} \times \det(\mathscr{A}_t^\theta)^{-1}\ . \tag{3.46}$$

Theorem 3.20. *Under the conditions of the lemma assume that the reduced Malliavin matrix $\mathscr{A}^\theta$ defined in* (2.50) *is invertible almost surely and $(\mathscr{A}^\theta)^{-1} \in L^p$ for some $p \geqslant 2$. Then the distribution of ξ_t is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$ with a continuous density in $\mathscr{C}_b(\mathbb{R}^d)$.*

Proof. We may apply Hölder's inequality to deduce

$$\mathbb{E}\big[\ |\det(\mathscr{D}_\theta \xi_t)^{-1}|^2\ \big] \leqslant \mathbb{E}\big[\ |\det(\nabla \xi_t)^{-1}|^{2\frac{p/2}{p/2-1}}\ \big] \times \mathbb{E}\big[\ |\det(\mathscr{A}_t^\theta)^{-1}|^p\ \big]\ .$$

This is finite. ∎

Remark 3.21. If one has (3.28) for any $p \geqslant 1$ it is possible to iterate the argument to conclude that the density is actually smooth ($\mathscr{C}^\infty$). We will not present a proof here but remark that the conclusion follows analogously to the Gaussian case in [Bel87, §3.3].

Chapter 4

The individually elliptic case. Bismut–Elworthy–Li formulae

We investigate gradient estimates of type (3.42) for the semigroup and the integration-by-parts formula in the case were the Gaussian respectively the Poissonian part is individually elliptic, i.e. the covariance operators are invertible. We can restrict our variational calculus to the Wiener respectively Poisson space individually. In the case of a continuous diffusion we are in the framework of the celebrated *Bismut–Elworthy–Li formula* [EL94]. It can be seen as a rudimentary version of an integration-by-parts formula. This chapter aims to make the reader familiar with the calculus as well as to give an overview over Bismut–Elworthy–Li type estimates in the literature.

4.1 Non-degenerate Gaussian: classical ellipticity

We assume that the diffusion coefficient σ is bounded and invertible with a bounded inverse. This requires that σ is square ($m = d$). Note however that for $m \geqslant d$ and σ of maximal rank we may define the (right) pseudo inverse

$$\sigma^{-1} := \sigma^*(\sigma\sigma^*)^{-1} \ .$$

We are in the classical framework of elliptic operators. Here we can restrict the variational approach to the Gaussian measure and consider the perturbation θ with

$$\begin{aligned} h_t &= \sigma(\xi_t)^{-1}\nabla\xi_t \ , \\ v(t,z) &\equiv 0 \ , \end{aligned} \qquad t \in [0,T], z \in \mathbb{R}^m \ . \tag{4.1}$$

Since σ is bounded and $\nabla\xi_t \in L^p(\Omega)$ for any $p \geqslant 1$ by Theorem 2.5 Condition 8 holds true and $\theta \in \Theta$. Recall that by Theorem 2.14 $\mathscr{D}_\theta\xi_t = \nabla\xi_t\mathscr{A}_t^\theta$, where in the present situation

$$\mathscr{A}_t^\theta = \int_0^t (\nabla\xi_{s-})^{-1}\sigma(\xi_{s-})\sigma(\xi_{s-})^{-1}(\nabla\xi_{s-})\mathrm{d}s = t \cdot \mathrm{Id} \ . \tag{4.2}$$

We see that with this choice of perturbation we reduced $\mathscr{A}_t^\theta$ to the quadratic variation of W and it suffices to take $\Psi \equiv t^{-1}$ deterministic. The integration-by-parts formula (3.29) with $f \in \mathscr{C}_b^1(\mathbb{R}^d)$ then simply reads

$$\mathbb{E}\big[\ \nabla f(\xi_t)\nabla\xi_t\ \big] = \mathbb{E}\big[\ f(\xi_t)\left(\int_0^t \big(\sigma(\xi_s)^{-1}\nabla\xi_s\big)^* \mathrm{d}W_s\right)\ \big]\ . \tag{4.3}$$

The Cauchy–Schwartz inequality together with Itô's isometry and the boundedness assumptions yield now

$$\begin{aligned} |\partial_y \mathcal{P}_t f(x)| = \frac{1}{t}|\mathbb{E}\big[\ \nabla f(\xi_t)\nabla\xi_t y\ \big]| &\leqslant \|f\|_\infty \frac{1}{t}\mathbb{E}\big[\ \int_0^t |\sigma(\xi_s)^{-1}\nabla\xi_s y|^2 \mathrm{d}s\ \big]^{\frac{1}{2}} \\ &\lesssim \|f\|_\infty \sup_{0\leqslant s\leqslant t}\mathbb{E}\big[\ |\nabla\xi_s y|^2\ \big]^{\frac{1}{2}} < \infty\ . \end{aligned} \tag{4.4}$$

We have already argued that this estimate allows to extend the gradient of the semigroup from $\mathscr{C}_b^1(\mathbb{R}^d)$ to $\mathscr{B}_b(\mathbb{R}^d)$. In particular we have verified the strong Feller property of $\mathcal{P}$.

Remark 4.1. An alternative approach consists in deriving a Bismut-Elworthy-Li formula in the non-degenerate Gaussian context conditionally on finitely many jumps and then take the limit. This is the strategy in [CF07] and does not require any Malliavin calculus. In particular compare (4.3) to formula (8) of [CF07, p.7]. Relying on the Malliavin calculus on Wiener space this strategy has also been followed in [WXZ15].

4.2 Non-degenerate Poisson noise

Assume that γ and all partial derivatives are bounded and that γ' is invertible with a bounded inverse. We mimic the procedure above in taking the perturbation θ to be

$$\begin{aligned} h_t &\equiv 0\ , \\ v(s,z) &= \gamma'(\xi_{s-},z)^{-1}(\mathrm{Id} + \nabla\gamma(\xi_{s-},z))\nabla\xi_{s-}|z|^2\ , \end{aligned} \qquad t\in[0,T], z\in\mathbb{R}^m\ . \tag{4.5}$$

It is readily verified that the perturbation satisfies Condition 8 (*cf.* Lemma 8) and hence $\theta \in \Theta_\varrho$. This strategy has been applied in [Nor88] for more general regularizing functions than $|z|^2$ in (4.5) (*cf.* the proof of Theorem 2.5) where the Gaussian part is absent completely, and more recently in [Tak10, p.582] where h is simultaneously taken according to (4.1). Again by Theorem 2.14 $\mathscr{D}_\theta\xi_t = \nabla\xi_t\mathscr{A}_t^\theta$, where we now have

$$\mathscr{A}_t^\theta = \mathrm{Id}\cdot \iint_0^t |z|^2\ N(\mathrm{d}z\mathrm{d}s)\ . \tag{4.6}$$

Here θ is such that $\mathscr{A}_t^\theta$ is the norm of the quadratic variation of the driving Lévy process Z. But since Z is a discontinuous martingale its quadratic variation is not deterministic and we have to take Ψ to be random. Indeed

$$\Psi = \left(\iint_0^t |z|^2\ N(\mathrm{d}z\mathrm{d}s)\right)^{-1}\ .$$

It can be shown similarly to the proof of Lemma 5.12 below that $\mathbb{E}[|\Psi|^p] < \infty$ for any $p > 0$. We need to calculate its Malliavin derivative. By formal differentiation we have

$$\frac{\partial}{\partial\lambda}\left(\Psi \circ \theta^\lambda\right) = \frac{\partial}{\partial\lambda}\left(\iint_0^t |z + v(s,z)\lambda|^2 \; N(\mathrm{d}z\mathrm{d}s)\right)^{-1} \tag{4.7}$$

$$= \frac{\iint_0^t 2(z + v(s,z)\lambda)^* v(s,z) \; N(\mathrm{d}z\mathrm{d}s)}{\left(\iint_0^t |z + v(s,z)\lambda|^2 \; N(\mathrm{d}z\mathrm{d}s)\right)^2} \;, \tag{4.8}$$

which for $\lambda = 0$ reduces to

$$\mathscr{D}_\theta \Psi = \frac{\iint_0^t 2z^* v(s,z) \; N(\mathrm{d}z\mathrm{d}s)}{\left(\iint_0^t |z|^2 \; N(\mathrm{d}z\mathrm{d}s)\right)^2} \;. \tag{4.9}$$

With the integration-by-parts formula (3.29) we obtain a Bismut-Elworthy-Li type formula that now reads

$$\mathbb{E}\big[\; \nabla f(\xi_t) \nabla \xi_t \;\big] = \mathbb{E}\big[f(\xi_t) \left\{ -\frac{\iint_0^t \varrho^{-1}\, \mathbf{div}_z \, \varrho v(s,z) \; \tilde{N}(\mathrm{d}z\mathrm{d}s)}{\iint_0^t |z|^2 \; N(\mathrm{d}z\mathrm{d}s)} + \frac{\iint_0^t 2z^* v(s,z) \; N(\mathrm{d}z\mathrm{d}s)}{\big(\iint_0^t |z|^2 \; N(\mathrm{d}z\mathrm{d}s)\big)^2} \right\} \big] \;.$$

Remark 4.2. A very similar formula has also been derived in [Tak10, Thm.1].

Chapter 5

Gradient estimates and densities. The jointly elliptic case

We now turn to the case where the covariance operators are only jointly elliptic. In this context neither $\sigma(x)$ nor $\gamma'(x)$ is invertible on its own and the choices of perturbation above are not feasible. Instead we apply techniques similar to the ones that are developed to handle Hörmander type conditions in a hypoelliptic framework. We will not consider such generalizations although in principle it is possible to go in this direction similarly to [Kun01, Kun11, Oh00] or [KT01].
Throughout this chapter we consider an *SDE*

$$\begin{cases} \mathrm{d}\xi_t = X(\xi_{t-}, \mathrm{d}t) \ , \\ \xi_0 = x \in \mathbb{R}^d \ , \end{cases} \tag{5.1}$$

on Ω with a semimartingale generator

$$X(x,t) = t\beta(x) + \sigma(x)A^{\frac{1}{2}}W_t + \iint_0^t \gamma(x,z)\tilde{N}(\mathrm{d}z\mathrm{d}s) \ , \tag{5.2}$$

where $A^{\frac{1}{2}}$ is a square root of a positive semidefinite covariance matrix A. The $\mathrm{d}z$ integral ranges over a suitable ball $\mathbb{B} \subset \mathbb{R}^m$ and the coefficients β, σ, γ are smooth and bounded vector- (resp. matrix-) functions with bounded derivatives of any order.

5.1 Assumptions

This section formulates the set of assumptions we need to impose in order to obtain regularity results for the *SDE*. The assumptions will be divided into properties of the coefficients, i.e. vector fields and the jump kernel, and into properties of the driving noise characterized by the Lévy triplet.

5.1.1 The Lévy triplet. Infinitesimal covariance and non degeneracy

We consider a Lévy triplet (b, A, ν) on $\mathbb{R}^m$. Let us introduce the following quantity

$$\sigma_\nu(\varepsilon) := \int_{|z|\leqslant\varepsilon} |z|^2 \nu(\mathrm{d}z) \ (< \infty) \ . \tag{5.3}$$

Condition 9 (Order condition)**.** We say that the Lévy measure ν satisfies an *order condition* if the function σ_ν is regularly varying at zero of order $\alpha \in (0, 2)$, i.e.

$$\liminf_{\varepsilon\searrow 0} \varepsilon^{-\alpha}\sigma_\nu(\varepsilon) \in (0, \infty) \quad \text{for some } \alpha \in (0, 2) \ . \tag{5.4}$$

An order condition allows us to define an infinitesimal covariance structure of the jumps that is comparable to the Gaussian covariance.

Definition 5.1 (Infinitesimal covariance)**.** For any $\varepsilon > 0$ define the $m \times m$ matrix

$$\Sigma_\nu(\varepsilon) := \frac{1}{\sigma_\nu(\varepsilon)} \int_{|z|<\varepsilon} z \otimes z \ \nu(\mathrm{d}z) \ . \tag{5.5}$$

The INFINITESIMAL COVARIANCE is defined as

$$A_\nu = \begin{cases} \liminf_{\varepsilon\searrow 0} \Sigma_\nu(\varepsilon) & , \text{ if } \nu \text{ satisfies the order condition,} \\ 0 & , \text{ otherwise.} \end{cases} \tag{5.6}$$

Here the limes inferior is taken with respect to the positive semidefinite partial order. This definition is not unique. However we can fix any representative.

Example 5.2. Let ν_α be the Lévy measure of an m-dimensional α-stable Lévy process for $0 < \alpha < 2$. Then there exists a (finite) spherical measure μ on $\mathbb{S}^{m-1}$ such that ν_α decomposes into a spherical and a radial measure (cf. [Sat99, Thm.14.3, p.77]), i.e.

$$\nu_\alpha(B) = \int_0^\infty \int_{\mathbb{S}^{m-1}} \mathbb{1}_B(r\theta) \ \mu(\mathrm{d}\theta) \frac{\mathrm{d}r}{r^{\alpha+1}} \ . \tag{5.7}$$

Hence ν_α satisfies the order condition 9 for α' given by $\alpha' = 2 - \alpha$. If μ is the uniform (Lebesgue) measure on $\mathbb{S}^{m-1}$, then ν_α is rotationally invariant. If μ is concentrated on the axes then Z is the sum of m independent scalar Lévy processes Z^j $(j = 1, \ldots, m)$ in each coordinate. In both cases however we have that

$$A_{\nu_\alpha} = \mathrm{Id} \in \mathbb{R}^{m\times m} \ . \tag{5.8}$$

This reflects the fact, that the coordinate jumps may be uncorrelated but not necessarily independent. To see that consider for $i \neq j$ and $\varepsilon > 0$

$$\begin{aligned} \int_{|z|<\varepsilon} z^i z^j \mathrm{d}\nu_\alpha &= \int_0^\varepsilon \int_0^{2\pi} r^2 \sin(\phi)\cos(\phi) \frac{\Gamma(m/2)\mathrm{d}\phi}{2\pi^{m/2}} \frac{\mathrm{d}r}{r^{\alpha+1}} \\ &= \frac{\varepsilon^{2-\alpha}}{2-\alpha} \frac{\Gamma(m/2)}{2\pi^{m/2}} \left(-\frac{1}{2}\cos^2(\phi)\right)\Big|_0^{2\pi} = 0 \ . \end{aligned}$$

Remark 5.3. We have seen in Example 5.2 that the order condition holds for any α-stable Lévy measure. We deduce that the order condition is also fulfilled by any Lévy measure ν that is bounded below by an α-stable measure $\nu_\alpha, \alpha \in (0,2)$, around the origin, i.e.

$$\nu(\{|z| \leqslant \varepsilon\}) \geqslant \nu_\alpha(\{|z| \leqslant \varepsilon\}) \ , \forall \varepsilon < \varepsilon_0 \ ,$$

for some $\varepsilon_0 > 0$ (*cf.* [IKM15]). In particular tempered stable measures fulfill an order condition.

As mentioned the infinitesimal covariance allows us to treat the Poisson random measure and the Gaussian component in a unified way. Indeed we define the non-degeneracy of the noise as follows.

Definition 5.4 (non-degeneracy of a Lévy process)**.** We say that the Lévy process on $\mathbb{R}^m$ with characteristic triplet (b, A, ν) is NON-DEGENERATE if the positive semidefinite $m \times m$ matrix $A + A_\nu$ is invertible, i.e.

$$\det(A + A_\nu) > 0 \ . \tag{5.9}$$

5.1.2 Ellipticity

Recall that γ' denotes the derivative of the jump kernel γ in the direction of z at $z = 0$. The ellipticity we require is stated as follows.

Condition 10 (Uniform ellipticity)**.** In the framework of (2.1), assume that the Lévy measure satisfies an order condition with $0 < \alpha < 2$. Denote

$$\Xi(x) := \sigma(x) A \sigma(x)^* + \gamma'(x) A_\nu \gamma'(x)^* \ . \tag{5.10}$$

We ask Ξ to be uniformly positive definite, i.e that there exists a constant $\kappa_0 > 0$ with

$$\langle \eta, \Xi(x)\eta \rangle \geqslant \kappa_0 |\eta|^2 \ , \forall x, \eta \in \mathbb{R}^d \ . \tag{5.11}$$

5.2 The Perturbation

We define a perturbation $\theta = (h, v)$, where

$$\begin{aligned} h_t &= \big((\nabla \xi_{t-})^{-1} \sigma(\xi_t) A^{\frac{1}{2}}\big)^* \ , \\ v(t,z) &= z \otimes z \big((\nabla \xi_{t-})^{-1} (\mathrm{Id} + \nabla \gamma(\xi_{t-}, z))^{-1} \gamma'(\xi_t, z)\big)^* \ , \end{aligned} \qquad t \in [0,T], z \in \mathbb{R}^m \ . \tag{5.12}$$

Let us also introduce the notation

$$v(t,z) = z \otimes z V(t,z)^* \text{ with } V(t,z) = \big((\nabla \xi_{t-})^{-1} (\mathrm{Id} + \nabla \gamma(\xi_{t-}, z))^{-1} \gamma'(\xi_t, z)\big) \ . \tag{5.13}$$

This choice would make $\mathscr{A}^\theta$ in (2.50) a positive semidefinite matrix. Let us verify that the choice is feasible. Since $\sigma, \gamma, \nabla\gamma$ and γ' are bounded we immediately see that for any $p \geqslant 1$ we have

$$\begin{aligned}\|h\|_{[0,T]} &\lesssim \|(\nabla\xi_{t-})^{-1}\|_{[0,T]} \in L^p(\Omega) \ , \\ \|v\|_{[0,T]} &\lesssim |z|^2\|(\nabla\xi_{t-})^{-1}\|_{[0,T]} \in L^p(\Omega \times \mathbb{R}^m, \mathbb{P} \times \nu) \ ,\end{aligned}$$

which verifies $\theta \in \Theta_p$. It remains to show that θ can be approximated by simple perturbations in Θ_0 and hence $\theta \in \Theta_\varrho$.
In order to do so we need another technical assumption.

Condition 11. We assume that the Lévy density ϱ satisfies

$$\int |z|^4 \frac{|\nabla\varrho(z)|^2}{\varrho(z)} \mathrm{d}z < \infty \ . \tag{5.14}$$

Remark 5.5. Condition 11 is a condition of the oscillation of ϱ at 0. The condition is satisfied whenever the quotient

$$\frac{|z||\nabla\varrho(z)|}{\varrho(z)} \tag{5.15}$$

is bounded. As usual in this chapter the integral is taken over any suitable environment of zero.

Remark 5.6. Condition 11 is satisfied by truncated stable and truncated stable-like Lévy measures. Indeed if $\varrho_\alpha(z) \sim |z|^{-(\alpha+1)}$ for some $\alpha \in (0, 2)$. Then the integral in (5.14) reads up to a constant

$$(\alpha+1)^2 \int |z|^4 \frac{|z|^{-2(\alpha+2)}}{|z|^{-(\alpha+1)}} \mathrm{d}z = (\alpha+1)^2 \int |z|^{1-\alpha} \mathrm{d}z < \infty \ .$$

5.2.1 Simple approximation

In order to apply any of our integration-by-parts settings (e.g. Theorem 3.10) we need to approximate h by $h_n \in \mathcal{H}_0$ and v by $v_n \in \mathcal{V}_0$. Note that $v = z \otimes zV(\omega, t, z)$ with $V = \big((\nabla\xi_{t-})^{-1}(\mathrm{Id} + \nabla\gamma(\xi_{t-}, z))^{-1}\gamma'(\xi_t, z)\big)^*$. Since $\nabla\gamma$ and γ' together with all other partial derivatives of γ are bounded, we see that V and $\nabla_z V$ are bounded above by a random variable $R_t \in L^p(\Omega, \mathbb{P})$ for any $p \geqslant 1$, $t \in [0, T]$. (a multiple of $\|(\nabla\xi)^{-1}\|_{[0,t]}$.) We denote further by $R_t^\star = \sup_{0 \leqslant s \leqslant t} R_t \in L^p$. Let us approximate v by $v_n \in \mathcal{V}_0, n \in \mathbb{N}$ the following way

$$v_n(\omega, t, z) := z \otimes z\chi_n(z)V(\omega, t, z)\mathbb{1}_{\{R_t^\star \leqslant n\}}(\omega, t) \in \mathcal{V}_0 \ , \tag{5.16}$$

where $\chi_n \geqslant 0$ is a smooth function that satisfies

$$\mathbb{1}_{\{\frac{1}{n} \leqslant |z| \leqslant n\}} \leqslant \chi_n(z) \leqslant \mathbb{1}_{\{\frac{1}{n+1} \leqslant |z| \leqslant n+1\}} \quad \text{with} \quad \frac{\nabla\chi_n}{1-\chi_n} \text{ bounded.} \tag{5.17}$$

Similarly we can approximate h. Since σ is bounded we may assume that also $\|h\|_{[0,t]} \leqslant R_t^\star$ almost surely and define

$$h_n(\omega,t) := h_t \mathbb{1}_{\{R_t^\star \leqslant n\}}(\omega,t) \in \mathcal{H}_0 \ . \tag{5.18}$$

We obviously have for all ω, t, z

$$[h - h_n](t) = \mathbb{1}_{\{R_t^\star > n\}} h(t) \ , \tag{5.19}$$

$$[v - v_n](t,z) = \big(\mathbb{1}_{\{R_t^\star > n\}} + (1-\chi_n(z))\mathbb{1}_{\{R_t^\star \leqslant n\}}\big) v(t,z) \ . \tag{5.20}$$

With this definition we have that $\theta_n := (h_n, v_n) \in \Theta_0$ provide a good approximating sequence for θ.

Lemma 5.7. *Assume that Condition 11 is satisfied. Then we have $\theta = (h,v) \in \Theta_\varrho \cap \Theta_p$ for any $p \geqslant 2$.*

Proof. We already observed that $\theta \in \Theta_p$. We show that v_n defined in (5.16) approximates v in $\mathcal{V}_\varrho$. We calculate the divergence of the matrix v with $V(z) = (V_{ij}(z))$. We obtain

$$\frac{\partial}{\partial z^j} v^{ji}(z) = \frac{\partial}{\partial z^j} \sum_k z^j z^k V_{ki}(z) = z^j V_{ji}(z) + \sum_k z^k V_{ki}(z) + \sum_k z^j z^k \frac{\partial}{\partial z^j} V_{ki}(z) \ .$$

Recall that the i-th entry of the vector $\mathbf{div}_z\, v$ is given by

$$\begin{aligned}
\mathrm{div}_z\, v^{\cdot i}(z) &= \sum_{j \leqslant m} \frac{\partial}{\partial z^j} v^{ji}(z) = \sum_j z^j V_{ji}(z) + m \sum_k z^k V_{ki}(z) + \sum_{jk} z^j z^k \frac{\partial}{\partial z^j} V_{ki}(z) \\
&= (m+1) \sum_k z^k V_{ki}(z) + \sum_k z^k z^* \nabla V_{ki}(z) \\
&= (m+1) z^* V_{\cdot i}(z) + z^* \nabla_z V_{\cdot i}(z) z \ .
\end{aligned}$$

Consequently we can estimate

$$\begin{aligned}
&|\varrho^{-1}\, \mathbf{div}_z\, ((v - v_n)\varrho)\,|^2 \\
&\qquad = |(1-\chi_n)\mathbb{1}_{\{R_t^\star \leqslant n\}}\big(\varrho^{-1}\, \mathbf{div}_z\big(v\varrho\big) + \frac{\nabla \chi_n}{1-\chi_n} v\big) + \mathbb{1}_{\{R_t^\star > n\}} \varrho^{-1}\, \mathbf{div}_z\big(v\varrho\big)|^2 \\
&\qquad \leqslant (1-\chi_n)^2 \left\{ |\varrho^{-1}\, \mathbf{div}_z\big(v\varrho\big)|^2 + |\frac{\nabla\chi_n}{1-\chi_n} v|^2 \right\} + \mathbb{1}_{\{R_t^\star > n\}} |\varrho^{-1}\, \mathbf{div}_z\big(v\varrho\big)|^2 \ ,
\end{aligned}$$

where we dropped the $\mathbb{1}_{\{R_t^\star \leqslant n\}}$ in the first term. With the use of the formula for $\mathbf{div}_z\, v$ the above line equals

$$\begin{aligned}
\big((1-\chi_n)^2 + \mathbb{1}_{\{R_t^\star > n\}}\big) &|(m+1) z^* V + z^* \nabla V z + (\nabla\varrho/\varrho) z \otimes z V|^2 \\
&+ (1-\chi_n)^2 \frac{|\nabla \chi_n|}{1-\chi_n} |z \otimes z V \log \rho|^2 \ .
\end{aligned}$$

Bearing in mind that $|\nabla\chi_n|(1-\chi_n)^{-1}$ is bounded and that $|V|$ and $|\nabla V|$ are bounded by $R_t^* \in L^p(\Omega)$ we finally estimate

$$
\begin{aligned}
\|v - v_n\|_{\mathcal{V}_\varrho} &= \mathbb{E}\Big[\iint_0^t |\varrho^{-1}\, \mathbf{div}_z\, ((v-v_n)\varrho)\, |^2 \mathrm{d}\nu \mathrm{d}s \Big] \\
&\leqslant \mathbb{E}\big[\, |R_t^\star|^2 \, \big] \iint_0^t (1-\chi_n)^2 \left((m+1)|z|^2 + |z|^4 + |z|^2 |\nabla\varrho|/\varrho\right)^2 \mathrm{d}\nu \mathrm{d}s \\
&\quad + \mathbb{E}\big[\, |R_t^\star|^2 \mathbb{1}_{\{R_t^\star > n\}} \, \big] \iint_0^t \left((m+1)|z|^2 + |z|^4 + |z|^2 |\nabla\varrho|/\varrho\right)^2 \mathrm{d}\nu \mathrm{d}s \\
&\quad + \mathbb{E}\big[\, |R_t^\star|^2 \, \big] \iint_0^t (1-\chi_n)^2 \Big(|z|^2 \log\rho\Big)^2 \mathrm{d}\nu \mathrm{d}s \\
&\to 0 \ ,
\end{aligned}
$$

by dominated convergence as n goes to infinity. Similarly the h_n approximate h in $\mathcal{H}$. ■

In order to safely apply the integration-by-parts framework of 3.1.3 we further need the following convergence results.

Theorem 5.8. *Let $\theta_n = (h_n, v_n) \in \Theta_0$ be defined by* (5.18) *and* (5.16)*. Then for any $p \geqslant 2$ we have*

$$
\mathscr{D}_{\theta_n}\xi_t \to \mathscr{D}_\theta \xi_t \text{ in } L^p \ . \tag{5.21}
$$

Proof. With the help of Theorem 2.13 we derive the *SDE* for the difference

$$
\begin{aligned}
\mathrm{d}[\mathscr{D}_\theta\xi_t - \mathscr{D}_{\theta_n}\xi_t] &= \nabla X(\xi_{t-}, \mathrm{d}t)[\mathscr{D}_\theta\xi_t - \mathscr{D}_{\theta_n}\xi_t] \\
&\quad + \sigma(\xi_{t-})[h - h_n](t)\mathrm{d}t + \int \gamma'(\xi_{t-}, z)[v - v_n](t,z) N(\mathrm{d}z\mathrm{d}t).
\end{aligned}
$$

Compensating the Poisson random measure we are able to apply the L^p-estimates of Proposition B.2. With the difference at $t = 0$ being zero we obtain

$$
\begin{aligned}
&\mathbb{E}|\mathscr{D}_\theta\xi_\cdot - \mathscr{D}_{\theta_n}\xi_\cdot|^p_{[0,t]} \lesssim_p \\
&\quad \mathbb{E}\Big[\Big(\big(\int_0^t |\nabla\beta(\xi_{s-})|\mathrm{d}s\big)^p + \iint |\nabla\gamma(\xi_{s-}, z)|^p \nu(\mathrm{d}z)\ \mathrm{d}s + \big(\int_0^t |\nabla\sigma(\xi_{s-})|^2 \mathrm{d}s\big)^{p/2} \\
&\quad + \big(\iint_0^t |\nabla\gamma(\xi_{s-}, z)|^2 \nu(dz)\mathrm{d}s\big)^{p/2}\Big)|\mathscr{D}_\theta\xi_\cdot - \mathscr{D}_{\theta_n}\xi_\cdot|^p_{[0,t]}\Big] \\
&\quad + \mathbb{E}\Big[\big(\int_0^t |\gamma'(\xi_{s-}, z)| \|v - v_n\|(s,z)\nu(\mathrm{d}z)\mathrm{d}s\big)^p + \int_0^t |\gamma'(\xi_{s-}, z)|^p \|v - v_n\|^p(s,z)\nu(\mathrm{d}z)\mathrm{d}s \\
&\quad \big(\int_0^t |\sigma(\xi_{s-})|^2 \|h - h_n\|^2(s)\mathrm{d}s\big)^{p/2} \ + \big(\int_0^t |\gamma'(\xi_{s-}, z)|^2 \|v - v_n\|^2(s,z)\nu(\mathrm{d}z)\mathrm{d}s\big)^{p/2}\Big].
\end{aligned}
$$

In view of (5.20) we observe that

$$
\iint_0^t |\gamma'(\xi_{s-}, z)|^p \|v\|^p (1 - \chi_n(z))^p \nu(\mathrm{d}z)\mathrm{d}s \searrow 0
$$

by dominated convergence. Also

$$\mathbb{E}\Big[\iint_0^t |\gamma'(\xi_{s-},z)|^p \|v\|^p \mathbb{1}_{\{R_t^\star \geqslant n\}} \nu(\mathrm{d}z)\mathrm{d}s \Big]$$
$$\leqslant \mathbb{E}\Big[(R_t^\star)^p \mathbb{1}_{\{R_t^\star \geqslant n\}} \iint_0^t |\gamma'(\xi_{s-},z)|^p |z|^{2p} \nu(\mathrm{d}z)\mathrm{d}s \Big] \searrow 0$$

∎

Remark 5.9. Applying the above theorem mutatis mutandis for graded *SDE* described in Section B.2 we also obtain that higher derivatives converge. In particular our assumptions imply

$$\mathscr{D}^2_{\theta_n}\xi_t \to \mathscr{D}^2_\theta \xi_t \text{ in } L^p \ . \tag{5.22}$$

Theorem 5.10. *For any $p \geqslant 2$ we have that $\mathscr{A}^\theta$ has an L^p-derivative in the direction of θ, furthermore we have*

$$\mathscr{A}_t^{\theta_n} \to \mathscr{A}_t^\theta,\ \mathscr{D}_{\theta_n}\mathscr{A}_t^{\theta_n} \to \mathscr{D}_\theta \mathscr{A}_t^\theta \ \textit{in } L^p \ . \tag{5.23}$$

Proof. Recall that by (2.50)

$$\begin{aligned}
\mathscr{A}_t^\theta - \mathscr{A}_t^{\theta_n} &= \int_0^t (\nabla\xi_{s-})^{-1}\sigma(\xi_{s-})(h-h_n)(s)\mathrm{d}s \\
&+ \iint_0^t (\nabla\xi_{s-})^{-1}(\mathrm{Id} + \nabla\gamma(\xi_{s-},z))^{-1}\gamma'(\xi_{s-},z)(v-v_n)(s,z)\ N(\mathrm{d}z\mathrm{d}s) \\
&= \mathbb{1}_{\{R_t^\star > n\}} \int_0^t (\nabla\xi_{s-})^{-1}\sigma(\xi_{s-})h_s\mathrm{d}s \\
&+ \iint_0^t \big[\mathbb{1}_{\{R_t^\star > n\}} + (1-\chi_n)(z)\big](\nabla\xi_{s-})^{-1}(\mathrm{Id} + \nabla\gamma(\xi_{s-},z))^{-1}\gamma'(\xi_{s-},z)v(s,z)\ N(\mathrm{d}z\mathrm{d}s) \ .
\end{aligned}$$

Using the L^p-estimates of Proposition B.2 one verifies similarly to the proof of Theorem 5.8 that $\mathscr{A}_t^{\theta_n} \to \mathscr{A}_t^\theta$ in L^p.
Now first observe that by the relation (2.49) we have $\mathscr{A}_t^\theta = (\nabla\xi_t)^{-1}\mathscr{D}_\theta\xi_t$ and hence its components are polynomials in the entries of $\mathscr{D}\xi$ and $\nabla\xi$. The existence of an L^p-derivative then follows from the chain rule (Proposition (2.3)).
To conclude that $\mathscr{D}_{\theta_n}\mathscr{A}_t^{\theta_n} \to \mathscr{D}_\theta\mathscr{A}_t^\theta$ in L^p we note that from (2.49) we can also deduce that

$$\mathscr{D}_\theta\mathscr{A}_t^\theta = (\nabla\xi_t)^{-1}\mathscr{D}^2_\theta\xi_t - (\nabla\xi_t)^{-1}(\mathscr{D}_\theta\nabla\xi_t)\mathscr{D}_\theta\xi_t \ . \tag{5.24}$$

Since all components on the right-hand side converge in L^p we conclude that also the left-hand side converges. ∎

5.3 Invertibility of the (reduced) Malliavin matrix

The crucial ingredient in deriving the semigroup estimates by means of the integration-by-parts formula is to show that the reduced Malliavin matrix $\mathscr{A}^\theta$ of (2.49) is invertible and the inverse is sufficiently integrable. We will establish the necessary results in this section. For the sake of readability we mostly suppress the superscript θ in this section. With the perturbation chosen according to (5.12) we have

$$\mathscr{A}_t^\theta = \int_0^t \mathscr{A}^1(s)\mathrm{d}s + \iint_0^t \mathscr{A}^2(s,z)\ N(\mathrm{d}z\mathrm{d}s)\ , \tag{5.25}$$

where

$$\begin{aligned}\mathscr{A}^1(t) &= (\nabla\xi_{t-})^{-1}\sigma(\xi_{t-})A\sigma(\xi_{t-})^*(\nabla\xi_{t-})^{-1*} \\ &= \left((\nabla\xi_{t-})^{-1}\sigma(\xi_{t-})a\right)^{\otimes 2}\ ,\end{aligned}\tag{5.26}$$

$$\begin{aligned}\mathscr{A}^2(t,z) &= (\nabla\xi_{t-})^{-1}(\mathrm{Id}+\nabla\gamma(\xi_{t-},z))^{-1}\gamma'(\xi_{t-},z) \\ &\qquad z\otimes z\gamma'(\xi_{t-},z)^*(\mathrm{Id}+\nabla\gamma(\xi_{t-},z))^{-1*}(\nabla\xi_{t-})^{-1*} \\ &= \left((\nabla\xi_{t-})^{-1}(\mathrm{Id}+\nabla\gamma(\xi_{t-},z))^{-1}\gamma'(\xi_{t-},z)z\right)^{\otimes 2}\ .\end{aligned}\tag{5.27}$$

$\mathscr{A}$ is clearly symmetric and positive semi-definite. The following theorem is the main result of this section.

Theorem 5.11. *Assume the uniform ellipticity condition 10. Then $\mathscr{A}_t^\theta$ and its inverse $(\mathscr{A}_t^\theta)^{-1}$ are L^p-differentiable in the direction of θ for any $p>0$. In particular both are in $L^p(\Omega)$.*

Proof. In view of the differentiability of $\mathscr{A}_t^\theta$ and Lemma 2.4 we need to prove that the determinant satisfies $\det(\mathscr{A}_t^\theta)^{-1}\in L^p(\Omega)$ for any $p>0$. This is established in Theorem 5.14 below. ■

The rest of this section is devoted to derive Theorem 5.14. The non-degeneracy of the positive semi-definite matrix can be established by investigating the positivity of the quadratic form $\langle\eta,\mathscr{A}_t^\theta\eta\rangle$ on vectors $\eta\in\mathbb{R}^d$. Indeed we have the following.

Lemma 5.12. *Under the uniform ellipticity condition 10 we have*

$$\sup_{|\eta|=1}\mathbb{E}\langle\eta,\mathscr{A}_t^\theta\eta\rangle^{-p}<\infty\ . \tag{5.28}$$

The proof is a direct consequence of the decay rate at $-\infty$ of the *moment generating function* (*Laplace transform*) of the random variable $\langle\eta,\mathscr{A}_t^\theta\eta\rangle$, i.e.

$$m_\eta(\lambda) := \mathbb{E}\left[\ \mathrm{e}^{\lambda\langle\eta,\mathscr{A}_t^\theta\eta\rangle}\ \right]\ , \tag{5.29}$$

defined for all $\lambda\in\mathbb{R}$ where the right hand side is finite, which in particular includes the negative half line.

Proof of Lemma 5.12. First observe that Condition 10 guarantees that $\langle \eta, \mathscr{A}_t^\theta \eta \rangle \geqslant 0$. A well known relation between negative moments of positive random variables and the moment generating function m_η (*cf.* [CDFG81]) states that

$$\mathbb{E}\big[\ \langle \eta, \mathscr{A}_t^\theta \eta \rangle^{-p}\ \big] = \frac{1}{\Gamma(p)} \int_0^\infty \lambda^p m_\eta(-\lambda) \mathrm{d}\lambda\ , \tag{5.30}$$

where the right hand side is uniformly bounded due to the uniform bound on m_v obtained in Lemma 5.13. ∎

Indeed we will prove the following estimate.

Lemma 5.13. *Assume the uniform ellipticity condition 10. For any $p \geqslant 1$ and $\lambda > 0$ there are constants $C_1, C_2 > 0$ independent of η and λ such that the following bound for m_η holds.*

$$m_\eta(-\lambda) \leqslant \mathbb{E}\Big[\exp\Big(- C_1 t \lambda^{-\frac{2-\alpha}{2}} \mathbb{E}\big[\ \|\nabla \xi\|_{[0,t]}^{-2}\ \big]\Big)\Big] + C_2 \lambda^{-p}\ . \tag{5.31}$$

The proof of the Lemma will be postponed to the end of this chapter. We will first state the important observation.

Theorem 5.14. *Under the uniform ellipticity condition 10 we have*

$$(\det \mathscr{A}_t^\theta)^{-1} \in L^p(\Omega)\ , \quad \textit{for all } p > 0\ .$$

Remark 5.15. Since $\mathbb{E}\big[\ \langle v, \mathscr{A}_t^\theta v \rangle^{-p}\ \big] < \infty$ we necessarily have $\mathbb{P}\big(\ \langle v, \mathscr{A}_t^\theta v \rangle^{-p} \geqslant x\ \big) = o(x^{-1})$ for large x. Hence we obtain

$$\mathbb{P}\left(\langle v, \mathscr{A}_t^\theta v \rangle \leqslant \varepsilon\right) = \mathbb{P}\left(\langle v, \mathscr{A}_t^\theta v \rangle^{-p} \geqslant \varepsilon^{-p}\right) \lesssim \varepsilon^p\ . \tag{5.32}$$

The claim that

$$(\det \mathscr{A}_t^\theta)^{-1} \in L^p(\Omega)\ , \quad \text{for all } p > 0\ ,$$

could then be deduced via an adaption of Lemma 2.3.1. in [Nua06, p.133].

Remark 5.16. We also remark that (5.32) can be deduced directly from Lemma 5.13 evoking de Bruijn's Tauberian theorem [BGT87, Theorem 4.12.9, p.254]. Such an argument has been proposed in the context of small deviations in [Sim04] and [AD09].

Instead we implement a direct proof based on a spectral argument.

Proof. We make use of the following formula. For any symmetric positive definite matrix $A \in \mathbb{R}^{d\times d}$ and any $p \in (0, \infty)$ we have

$$\det(A)^{-p} \leqslant \frac{1}{\Gamma(p)^d} \int_0^\infty \int_{\mathbb{S}^{d-1}} r^{d(2p-1)}\, \mathrm{e}^{-r^2 \langle \bar{x}, A\bar{x} \rangle}\, \mathrm{d}\bar{x} \mathrm{d}r \tag{5.33}$$

(actually we have "$\simeq$", see e.g. [BGJ87, Lemma 7-29, p.92]). If we evoke the statement for $\mathscr{A}$ together with the uniform estimate of Lemma 5.13 for some $p' > p$ we obtain

$$\begin{aligned}\mathbb{E}\big[\ \det(\mathscr{A}_t^\theta)^{-p}\ \big] &\leqslant \frac{1}{\Gamma(p)^d}\int_0^\infty \int_{\mathbb{S}^{d-1}} r^{d(2p-1)}\mathbb{E}\big[\ \mathrm{e}^{-r^2\langle \bar{x}, \mathscr{A}_t^\theta \bar{x}\rangle}\ \big]\mathrm{d}\bar{x}\mathrm{d}r \\ &\leqslant \frac{\mathrm{Vol}(\mathbb{S}^{d-1})}{\Gamma(p)^d}\Big\{1 + \int_1^\infty r^{d(2p-1)}\Big\{\mathrm{e}^{-C_1 r^{-2+\alpha}} + C_2 r^{-2p'}\Big\}\,\mathrm{d}r\Big\} \\ &< \infty\ ,\end{aligned}$$

for two appropriately chosen constants $C_1, C_2 > 0$. ■

To continue with the proof of Lemma 5.13 we need an explicit expression for m_η.

Lemma 5.17. *Fix a test direction $\eta \in \mathbb{S}^{d-1}$. The* moment generating function *(*Laplace transform*) of $\langle \eta, \mathscr{A}\eta\rangle$ is defined on the negative half-line and for any $\lambda \geqslant 0$ we have*

$$\begin{aligned}m_\eta(-\lambda) = \mathbb{E}\exp\Big(-\lambda\int_0^t \langle \eta, \mathscr{A}_s^1\eta\rangle \mathrm{d}s \\ + \iint_0^t (\mathrm{e}^{-\lambda\langle \eta, \mathscr{A}_{s,z}^2\eta\rangle} - 1)\ \nu(\mathrm{d}z)\mathrm{d}s\Big)\ .\end{aligned} \tag{5.34}$$

Proof. Since $\mathscr{A}^\theta$ solves (5.25) it follows by linearity that $Y := -\lambda\langle \eta, \mathscr{A}^\theta\eta\rangle$ satisfies

$$\mathrm{d}Y_t = f(s)\mathrm{d}s + \int g(s,z)N(\mathrm{d}z\mathrm{d}s)\ ,$$

with $f(s) = -\lambda\langle \eta, \mathscr{A}_s^1\eta\rangle$, $g(s,z) = -\lambda\langle \eta, \mathscr{A}_{s,z}^2\eta\rangle$. By Itô's formula

$$\begin{aligned}\mathrm{d}\left[\mathrm{e}^Y\right]_t &= \mathrm{e}^{Y_{t-}} f(t)\mathrm{d}t + \mathrm{e}^{Y_{t-}}\int \big[\mathrm{e}^g - 1\big](t,z)\ N(\mathrm{d}z\mathrm{d}t) \\ &= \mathrm{e}^{Y_{t-}} f(t)\mathrm{d}t + \mathrm{e}^{Y_{t-}}\int \big[\mathrm{e}^g - 1\big](t,z)\ \nu(\mathrm{d}z)\mathrm{d}t \\ &\quad + \mathrm{e}^{Y_{t-}}\int \big[\mathrm{e}^g - 1\big](t,z)\ \tilde{N}(\mathrm{d}z\mathrm{d}t)\ .\end{aligned}$$

Taking expectation and integrating under the expectation yields

$$\begin{aligned}\mathbb{E}\big[\ \mathrm{e}^{Y_t}\ \big] &= \mathbb{E}\big[\ 1 + \int_0^t \mathrm{e}^{Y_{s-}}\big\{f(s) + \int \big[\mathrm{e}^g - 1\big](s,z)\ \nu(\mathrm{d}z)\big\}\mathrm{d}t\ \big] \\ &= \mathbb{E}\big[\ \exp\big\{\int_0^t f(s)\mathrm{d}s + \iint_0^t \big[\mathrm{e}^g - 1\big](s,z)\ \nu(\mathrm{d}z)\mathrm{d}s\big\}\ \big]\ ,\end{aligned}$$

which is the desired result. ■

We will see that the smoothing property expressed by the decay of m_η is produced by a small jump size limit. We therefore introduce a different integrand that describes the small jump size limit in view of Condition 10. We define

$$\mathscr{A}^0(t,z) = (\nabla\xi_{t-})^{-1}\gamma'(\xi_{t-})z \otimes z\gamma'(\xi_{t-})^*(\nabla\xi_{t-})^{-1*} = \big((\nabla\xi_{t-})^{-1}\gamma'(\xi_{t-})z\big)^{\otimes 2} \ . \tag{5.35}$$

This is an approximation of $\mathscr{A}^1$ for small $|z|$. In deed we have the following.

Lemma 5.18. *By the boundedness of γ' and $\nabla\gamma$ we have*

$$\|\mathscr{A}^2(t,z) - \mathscr{A}^0(t,z)\| \lesssim \|(\nabla\xi)_{t-}^{-1}\|^2|z|^3$$

uniformly over all z in a ball around zero.

Proof. Note that

$$\begin{aligned}
&|\gamma'(x,z)zz^*\gamma'(x,z)^* - \gamma'(x,0)zz^*\gamma'(x,0)^*| \\
&= |\gamma'(x,z)zz^*\big(\gamma'(x,z)-\gamma'(x,0)\big)^* + \Big(\gamma'(x,z)-\gamma'(x,0)\Big)zz^*\gamma'(x,0)^*| \\
&\leqslant \big(\|\gamma'(x,0)\| + \|\gamma'(x,z)\|\big)|z|^2\|\gamma'(x,z)-\gamma'(x,0)\| \lesssim |z|^3
\end{aligned}$$

since all partial derivatives of γ are bounded and z is in a compactum. Finally we recall that $\operatorname{supp}\nu$ is small enough such that $\|(\mathrm{Id}+\nabla\gamma(x,z))^{-1}\|$ is bounded. ■

We seek to apply a trick similar to [IK06, Kun11, Ish13]. We define two random quantities

$$\widetilde{\mathscr{A}_\delta}(\eta) = \int\limits_{[0,t]} \langle\eta, \mathscr{A}_s^1\eta\rangle \mathrm{d}s + \frac{1}{\sigma_\nu(\delta)} \iint\limits_{[0,t]\times\{|z|<\delta\}} \langle\eta, \mathscr{A}_{s,z}^2\eta\rangle \wedge \delta^\beta \ \nu(\mathrm{d}z)\mathrm{d}s \ , \tag{5.36}$$

$$\widehat{\mathscr{A}_\delta}(\eta) = \int\limits_{[0,t]} \langle\eta, \mathscr{A}_s^1\eta\rangle \mathrm{d}s + \frac{1}{\sigma_\nu(\delta)} \iint\limits_{[0,t]\times\{|z|<\delta\}} \langle\eta, \mathscr{A}_{s,z}^0\eta\rangle \ \nu(\mathrm{d}z)\mathrm{d}s \ . \tag{5.37}$$

Recall that σ_ν defined in (5.3) converges to zero as $\delta \searrow 0$. The main task is to ensure the equivalence of the two asymptotically for small δ. The latter quantity it is inverted easily.

Lemma 5.19. *Under the uniform ellipticity condition 10*

$$\mathbb{E}\big[\ \widehat{\mathscr{A}_\delta}(\eta)^{-p}\ \big] \leqslant \big(\kappa_0 t\big)^{-p}\mathbb{E}\big[\ \|\nabla\xi\|_{[0,t]}^{2p}\ \big] < \infty \ .$$

Proof. It is readily seen that for any $0 < \delta \leqslant 1$

$$\begin{aligned}
\widehat{\mathscr{A}_\delta}(\eta) &\geqslant \int_0^t \langle(\nabla\xi_s)^{-1*}\eta, \Xi(\xi_s)(\nabla\xi_s)^{-1*}\eta\rangle\mathrm{d}s \\
&\geqslant \int_0^t \kappa_0|\eta^*(\nabla\xi_s)^{-1}|^2\mathrm{d}s \\
&\geqslant \int_0^t \kappa_0 \inf_{|y|=1}|(\nabla\xi_s)^{-1}y|^2\mathrm{d}s \\
&= \int_0^t \kappa_0\|\nabla\xi_s\|^{-2}\mathrm{d}s \\
&\geqslant \kappa_0 t\|\nabla\xi\|_{[0,t]}^{-2} \ .
\end{aligned} \tag{5.38}$$

Where we used the fact that the absolute value of left and right eigenvalues coincide and that the eigenvalues of a matrix coincide with the reciprocals of the eigenvalues of the inverse matrix. Taking the expectation of the inverse we conclude for $p \geqslant 1$

$$\mathbb{E}\big[\ (\widehat{\mathscr{A}_\delta}(v))^{-p}\ \big] \leqslant \big(\kappa_0 t\big)^{-p}\mathbb{E}\big[\ \|\nabla\xi\|_{[0,t]}^{2p}\ \big]\ .$$

■

It is not possible to obtain a similar estimate for $\widetilde{\mathscr{A}_\delta}(v)$ for fixed $\delta > 0$. Nevertheless we obtain a control for δ "small enough". In fact define the random variable

$$\delta^* := \inf\Big\{0 < \varrho \leqslant 1 : |\widetilde{\mathscr{A}_\delta} - \widehat{\mathscr{A}_\delta}| \geqslant \frac{1}{2}\widehat{\mathscr{A}_\delta}\ ,\quad \forall 0 < \delta \leqslant \varrho\Big\}\ . \tag{5.39}$$

The following Lemma is an adaption of [Kun11, Lemma 2.4, p.18] (see also [Ish13, pp.157-160]).

Lemma 5.20. *We have for any $p \geqslant 1$*

$$\mathbb{P}\big(\ \delta^* < \delta\ \big) \lesssim_p \delta^p \qquad \textit{for any } \delta \geqslant 0\ . \tag{5.40}$$

Proof. First we fix β with $\alpha < \beta < 2$. Denote $E(\delta) = \big\{\omega : \sup_{s<t,|z|\leqslant\delta}\langle\eta, \mathscr{A}^2_{s,z}(\omega)\eta\rangle \leqslant \delta^\beta\big\}$. Then we have on $E(\delta)$

$$|\langle\eta, \mathscr{A}^2_{t,z}\eta\rangle \wedge \delta^\beta - \langle\eta, \mathscr{A}^0_{t,z}\eta\rangle| \lesssim \|(\nabla\xi)^{-1}_{t-}\|^2|z|^3\ .$$

Hence we have for $q > 0$

$$\begin{aligned}
&\mathbb{E}\Big[\ \Big|\frac{1}{\sigma_\nu(\delta)}\iint\limits_{\substack{0<s<t\\|z|<\delta}}\langle\eta, \mathscr{A}^2_{s,z}\eta\rangle \wedge \delta^\beta - \langle\eta, \mathscr{A}^0_{s,z}\eta\rangle)\nu(\mathrm{d}z)\mathrm{d}s\Big|^q \mathbb{1}_{E(\delta)}\ \Big]\\
&\lesssim \frac{\delta^q t^q}{\sigma_\nu(\delta)^q}\Big(\int\limits_{|z|<\delta}|z|^2\nu(\mathrm{d}z)\Big)^q\mathbb{E}\big[\ \|\sup_{0<s\leqslant t}(\nabla\xi)^{-1}_{s-}\|^{2q}\ \big]\\
&\lesssim \delta\sup_{0<s\leqslant t}\mathbb{E}\big[\ \|(\nabla\xi_s)^{-1}\|^{2q}\ \big]\ .
\end{aligned}$$

Clearly this is also an upper bound for $\mathbb{E}\big[\ |\widetilde{\mathscr{A}_\delta} - \widehat{\mathscr{A}_\delta}|^q\mathbb{1}_{E(\delta)}\ \big]$. On the other hand we see that $E(\delta)^c = \big\{\sup_{s<t,|z|\leqslant\delta}\langle\eta, \mathscr{A}^2_{s,z}\eta\rangle > \delta^\beta\big\}$ and Chebychev's inequality yields

$$\mathbb{P}\big(\ E(\delta)^c\ \big) \leqslant \delta^{-\beta q}\mathbb{E}\big[\sup_{s<t,|z|\leqslant\delta}\langle\eta, \mathscr{A}^2_{s,z}\eta\rangle^q\ \big] \leqslant \delta^{(2-\beta)q}\sup_{s<t}\mathbb{E}\big[\ \|(\nabla\xi_t)^{-1}\|^{2q}\ \big]\ .$$

We can now estimate

$$\begin{aligned}
\mathbb{P}\big(\ \delta^* < \delta\ \big) &\leqslant \mathbb{P}\Big(\ \frac{|\widetilde{\mathscr{A}_\delta} - \widehat{\mathscr{A}_\delta}|}{\widehat{\mathscr{A}_\delta}} \geqslant \frac{1}{2}\ \Big)\\
&\leqslant \mathbb{P}\Big(\ \Big\{\frac{|\widetilde{\mathscr{A}_\delta} - \widehat{\mathscr{A}_\delta}|}{\widehat{\mathscr{A}_\delta}} \geqslant \frac{1}{2}\Big\} \cap E(\delta)\ \Big) + \mathbb{P}\big(\ E(\delta)^c\ \big)\\
&\leqslant 2^q\mathbb{E}\big[\ |\widetilde{\mathscr{A}_\delta} - \widehat{\mathscr{A}_\delta}|^{2q}\mathbb{1}_{E(\delta)}\ \big]^{1/2} \times \mathbb{E}\big[\ \widehat{\mathscr{A}_\delta}^{-2q}\ \big]^{1/2} + \mathbb{P}\big(\ E(\delta)^c\ \big)\ ,
\end{aligned}$$

where the last line is again an application of the Chebychev- and Cauchy-Schwartz inequalities. Since the second expectation is bounded uniformly in δ by Lemma 5.19 we see that there is a constant c independent of δ such that

$$\mathbb{P}\big(\ \delta^* < \delta\ \big) \leqslant c\left(\delta^{(3-\alpha)q} + \delta^{(2-\beta)q}\right) = O\big(\delta^{(2-\beta)q}\big)\ .$$

The claim follows for $q = \frac{p}{2-\beta}$. ■

Finally we are able do deliver the proof of Lemma 5.13.

Proof of Lemma 5.13. The moment generating function m_η is given in Lemma 5.17. We first observe that the local integrand in the exponent of (5.29) is of the form

$$\begin{aligned}\lambda\langle\eta, \mathscr{A}^1_s\eta\rangle &= \lambda\langle(\nabla\xi_{s-})^{-1*}\eta, \sigma(\xi_{s-})A\sigma(\xi_{s-})^*(\nabla\xi_{s-})^{-1*}\eta\rangle \\ &= \lambda\langle V_s, \sigma(\xi_{s-})A\sigma(\xi_{s-})^*V_s\rangle\ , \quad \lambda > 0\ .\end{aligned}$$

For a previsible process $V_s = (\nabla\xi_{s-})^{-1*}\eta$ taking values in $\mathbb{R}^d$. We investigate the non-local integral in the exponent for large λ.

$$\mathrm{e}^{\lambda\langle\eta,\mathscr{A}^2_{s,z}\eta\rangle} - 1 \geqslant \frac{1}{2}\big(\lambda\langle\eta, \mathscr{A}^2_{s,z}\eta\rangle \wedge 1\big) = \frac{\lambda}{2}\big(\langle\eta, \mathscr{A}^2_{s,z}\eta\rangle \wedge \lambda^{-1}\big)\ .$$

We fix some $\alpha < \beta < 2$. We may assume that λ is large enough such that $\frac{1}{2}\sigma_\nu(\lambda^{-\frac{1}{\beta}}) \leqslant 1$. We can then estimate the exponent of (5.29) by

$$\begin{aligned}&\lambda\int_0^t \langle\eta, \mathscr{A}^1_s\eta\rangle\mathrm{d}s + \iint_0^t (\mathrm{e}^{\lambda\langle\eta,\mathscr{A}^2_{s,z}\eta\rangle} - 1)\ \nu(\mathrm{d}z)\mathrm{d}s \\ &\geqslant \lambda\int_0^t \langle\eta, \mathscr{A}^1_s\eta\rangle\mathrm{d}s + \frac{\lambda}{2}\iint_0^t \langle\eta, \mathscr{A}^2_{s,z}\eta\rangle \wedge \lambda^{-1}\ \nu(\mathrm{d}z)\mathrm{d}s \\ &\geqslant \frac{\lambda}{2}\sigma_\nu(\lambda^{-\frac{1}{\beta}})\Big(\int_0^t \langle\eta, \mathscr{A}^1_s\eta\rangle\mathrm{d}s + \iint\limits_{[0,t]\times\left\{|z|\leqslant\lambda^{-1/\beta}\right\}} \frac{\langle\eta, \mathscr{A}^2_{s,z}\eta\rangle \wedge \lambda^{-1}}{\sigma_\nu(\lambda^{-\frac{1}{\beta}})}\ \nu(\mathrm{d}z)\mathrm{d}s\Big) \\ &= \frac{1}{2}\delta^{-\beta}\sigma_\nu(\delta) \times \widetilde{\mathscr{A}_\delta}\ , \text{ with } \delta = \lambda^{-\frac{1}{\beta}}\ .\end{aligned}$$

Recall the definition of δ^* to see that for sufficiently small δ (large λ such that $\delta^* \geqslant \delta$) the terms are further bounded below by

$$\begin{aligned}\frac{1}{4}\delta^{-\beta}\sigma_\nu(\delta) \times \widehat{\mathscr{A}_\delta} &\geqslant \frac{1}{4}\delta^{-\beta}\sigma_\nu(\delta)\int_0^t \langle(\nabla\xi_s)^{-1*}\eta, \Xi(\xi_s)(\nabla\xi_s)^{-1*}\eta\rangle\mathrm{d}s \\ &\geqslant \frac{1}{4}\delta^{-\beta}\sigma_\nu(\delta)\kappa_0\int_0^t |\eta^*(\nabla\xi_s)^{-1}|^2\mathrm{d}s \\ &\geqslant \frac{1}{4}\delta^{-\beta}\sigma_\nu(\delta)\kappa_0 t\|\nabla\xi\|^{-2}_{[0,t]}\ ,\end{aligned}$$

as in the proof of Lemma 5.19. Hence we obtain for $0 < \delta \leqslant 1$ with Jensen's inequality

$$m_\eta(-\lambda) \leqslant \mathbb{E}\Big[\exp\Big(-\frac{1}{4}\delta^{-\beta}\sigma_\nu(\delta) \times \widehat{\mathscr{A}_\delta}\Big)\mathbb{1}_{\{\lambda>1/\delta^*\}}\Big] + \mathbb{P}\Big(\delta^* < 1/\lambda\Big) \tag{5.41}$$

$$\leqslant \exp\Big(-\frac{1}{4}\delta^{-\beta}\sigma_\nu(\delta)\mathbb{E}\big[\ \widehat{\mathscr{A}_\delta}\ \big]\Big) + \mathbb{P}\Big(\delta^* < 1/\lambda\Big) \tag{5.42}$$

$$\leqslant \exp\Big(-\frac{1}{4}\delta^{-\beta}\sigma_\nu(\delta)\kappa_0 t\mathbb{E}\big[\ \|\nabla\xi\|_{[0,t]}^{-2}\ \big]\Big) + \mathbb{P}\Big(\delta^* < 1/\lambda\Big)\ , \tag{5.43}$$

where $\mathbb{E}\big[\ \|\nabla\xi\|_{[0,t]}^{-2}\ \big] > 0$. By Lemma 5.20 the claim is proven. ∎

Chapter 6

Jump diffusions on submanifolds in $\mathbb{R}^d$

In this chapter we investigate *SDE* where the state space is a manifold $\mathbb{M}$. Formally we write

$$\begin{cases} \mathrm{d}\xi_t = X(\xi_{t-}, \mathrm{d}t) \ , \\ \xi_0 = x \in \mathbb{M} \ , \end{cases} \tag{6.1}$$

where again

$$X(x,t) = t\beta(x) + \sigma(x)A^{\frac{1}{2}}W_t + \iint_0^t \gamma(x,z)\tilde{N}(\mathrm{d}z\mathrm{d}s) \ . \tag{6.2}$$

We assume that the coefficients of X are sufficiently smooth and such that the coefficients as well all partial derivatives are bounded. Then there exists a unique strong solution for any initial condition (at least in an ambient $\mathbb{R}^d$). This will be the standing assumption in this chapter. Again without loss of generality we may assume that $A^{\frac{1}{2}} = \mathrm{diag}(a_1, \ldots, a_m)$.

6.1 Invariant submanifolds of codimension 1

Assume that we have a submanifold $\mathbb{M} \subset \mathbb{R}^d$ that we take for simplicity of codimension 1 given as the zero set of a smooth function $\psi \in \mathscr{C}^\infty(\mathbb{R}^d)$, i.e.

$$\mathbb{M} = \psi^{-1}(\{0\}) \subset \mathbb{R}^d \ . \tag{6.3}$$

We want to derive conditions that ensure that the flow ξ generated by a jump diffusion of type (6.1) leaves $\mathbb{M}$ invariant. First note that if $\mathbb{M}$ is given by (6.3) then its tangent bundle $T\mathbb{M}$ is characterized as the orthogonal complement of the gradient of ψ, i.e.

$$y \in T_x\mathbb{M}, x \in \mathbb{M} \iff \langle \nabla\psi(x), y\rangle = 0 \ . \tag{6.4}$$

Example 6.1. The simplest example and the one we are specifically interested in is the $d-1$-dimensional unit sphere $\mathbb{S}^{d-1}$ which is trivially given as the level set of the norm, hence we may choose

$$\psi(x) = |x|^2 - 1 \ , \quad x \in \mathbb{R}^d \ . \tag{6.5}$$

We make the following observation. Consider a smooth vector field $X : \mathbb{M} \to T\mathbb{M}$. Since by (6.4) $\langle \nabla\psi(x), X(x)\rangle$ vanishes constantly on $\mathbb{M}$ its derivative in tangent direction vanishes also. Therefore we have

$$\nabla\langle \nabla\psi(x), X(x)\rangle y = 0 \ , \quad y \in T_x\mathbb{M}, x \in \mathbb{M} \ . \tag{6.6}$$

Remark 6.2. Of course the analysis extends directly to submanifolds of higher codimension if they are given as a common zero set of a finite set of smooth functions $\{\psi_1, \dots, \psi_k\}$, that is

$$\mathbb{M} = \cap_{i=1,\dots,k}\mathbb{M}_i, \text{ with } \mathbb{M}_i = \psi_i^{-1}(\{0\}) \ . \tag{6.7}$$

Now consider the solution ξ to (6.1). Let us first clarify what we mean if we say that ξ leaves $\mathbb{M}$ invariant.

Definition 6.3. A Borel subset B of $\mathbb{R}^d$ is said to be INVARIANT for the stochastic flow ξ or in other words ξ *leaves* B *invariant* if

$$\mathbb{P}\big(\ \xi_t \in B,\ \forall t > 0\ \big) = 1 \tag{6.8}$$

whenever $\mathbb{P}\big(\ \xi_0 \in B\ \big) = 1$.

The main result of this chapter is the following invariant manifold theorem.

Theorem 6.4. *Assume that*

(i) $\nabla\psi \cdot \sigma \equiv 0 \in \mathbb{R}^m$ *on* $\mathbb{M}$ *(*$\sigma \in T\mathbb{M}^{\otimes m}$*)*

(ii) $\psi \circ (\cdot + \gamma(\cdot, z)) \equiv 0$ *on* $\mathbb{M}$ *(*$x + \gamma(x, z) \in \mathbb{M}$*) for* ν*–almost every* $z \in \mathbb{R}^m$.

(iii) Assume that

$$\beta(x) - \int \big(\gamma(x, z) - \gamma'(x)z\big)\ \nu(\mathrm{d}z) - \frac{1}{2}\sum a_j^2 \nabla\sigma_{\cdot j}(x)\sigma_{\cdot j}(x) \in T_x\mathbb{M} \ . \tag{6.9}$$

Then ξ *leaves* $\mathbb{M}$ *invariant.*

Before giving a proof let us investigate the meaning of (6.9) closer. We recall that γ' denotes $\nabla_z\gamma(x, z)|_{z=0}$. We observe that $x + \gamma(x, z) \in \mathbb{M}$ for all $x \in \mathbb{M}$ and z. Letting $z \in \operatorname{supp}(\nu)$ tend to zero our assumptions imply that also $\nabla\psi(x)\gamma'(x)z = 0$, i.e. $\gamma'z$ is necessarily tangent (at least if $\operatorname{supp}\nu$ contains an environment of 0). Hence, in order for the integral

$$\int \big(\gamma(x, z) - \gamma'(x)z\big)\ \nu(\mathrm{d}z)$$

to be well defined, it is necessary that the variation in non-tangent directions is finite.
The following corollary is formulated in the spirit of [FTT14] (where it is stated in infinite dimensions in order to model interest rate curves in a financial application).

Corollary 6.5. *Let n be a normal vector field on $\mathbb{M}$ (unique up to the sign). Assume that $\int |\langle n(x), \gamma(x,z)\rangle| \nu(\mathrm{d}z) < \infty$ for any $x \in \mathbb{M}$. Then, in the situation of the theorem equation* (6.9) *is equivalent to*

$$\beta(x) + \int \langle n(x), \gamma(x,z)\rangle n(x)\nu(\mathrm{d}z) - \frac{1}{2}\sum a_j^2 \nabla\sigma_{\cdot j}(x)\sigma_{\cdot j}(x) \in T_x\mathbb{M} \, . \tag{6.10}$$

Proof of the theorem. If Z has finite variation (i.e. $\int_{|z|\leqslant\varepsilon} |z|\nu(\mathrm{d}z) < \infty$ for some (all) $\varepsilon > 0$ and $A = 0$) this is easily seen since ordinary differentiation yields

$$\mathrm{d}\psi(\xi_t(x)) = \nabla\psi(\xi_t(x))\beta(\xi_t(x))\mathrm{d}t + \int \psi\big(\xi_{t-}(x) + \gamma(\xi_{t-}(x), z)\big) - \psi\big(\xi_{t-}(x)\big) \; N(\mathrm{d}z\mathrm{d}t) \equiv 0 \, ,$$

almost surely which ensures that $\psi(\xi_t(x)) \equiv 0$ almost surely whenever $x \in \mathbb{M}$.

In the case of infinite variation we consider $A \succcurlyeq 0$ and we may set that $A = \mathrm{Id}$. Further $\int_{|z|\leqslant\varepsilon} |z|\nu(\mathrm{d}z) = \infty$. We observed that necessarily $\langle \nabla\psi(x), \gamma'(x)z\rangle = 0$. We apply Itô's formula to obtain

$$\begin{aligned}
\mathrm{d}\psi(\xi_t(x)) = {} & \nabla\psi(\xi_t(x))\beta(\xi_t)\mathrm{d}t + \nabla\psi(\xi_t(x))\sigma(\xi_t)\mathrm{d}W_t \\
& + \frac{1}{2}\operatorname{Tr}\left\{\sigma(\xi_t(x))^*\nabla^2\psi(\xi_t(x))\sigma(\xi_t(x))\right\}\mathrm{d}t \\
& + \int \psi\big(\xi_{t-}(x) + \gamma(\xi_{t-}(x), z)\big) - \psi\big(\xi_{t-}(x)\big) \; \tilde{N}(\mathrm{d}z\mathrm{d}t) \\
& + \int \psi\big(\xi_{t-}(x) + \gamma(\xi_{t-}(x), z)\big) - \psi\big(\xi_{t-}(x)\big) - \nabla\psi(\xi_{t-}(x))\gamma(\xi_{t-}(x), z) \; \nu(\mathrm{d}z)\mathrm{d}t \, .
\end{aligned}$$

The assumptions assert that the martingale terms vanish in $L^2(\Omega)$ and a.s. This follows from Itô's isometry since,

$$\begin{aligned}
& \mathbb{E}\Big[\; \Big|\int_0^t \nabla\psi(\xi_s)\sigma(\xi_s)\mathrm{d}W_s\Big|^2 + \Big|\iint_0^t \psi\big(\xi_{s-} + \gamma(\xi_{s-}, z)\big) - \psi\big(\xi_{s-}\big) \; \tilde{N}(\mathrm{d}z\mathrm{d}s)\Big|^2 \; \Big] \\
= {} & \mathbb{E}\Big[\; \int_0^t |\nabla\psi(\xi_s)\sigma(\xi_s)|^2\mathrm{d}s + \iint_0^t |\psi\big(\xi_{s-} + \gamma(\xi_{s-}, z)\big) - \psi\big(\xi_{s-}\big)|^2 \; \nu(\mathrm{d}z)\mathrm{d}s \; \Big] = 0 \, .
\end{aligned}$$

By Doob's maximal inequality this extends to the whole trajectory up to t. We conclude that $\psi(\xi)$ is of finite variation and

$$\begin{aligned}
\mathrm{d}\psi(\xi_t(x)) &= \nabla\psi(\xi_t(x))\Big(\beta(\xi_t) - \int \gamma(\xi_{t-}(x),z) - \gamma'(\xi_{t-}(x))z\ \nu(\mathrm{d}z)\Big)\mathrm{d}t \\
&\quad + \frac{1}{2}\operatorname{Tr}\left\{\sigma(\xi_t(x))^*\nabla^2\psi(\xi_t(x))\sigma(\xi_t(x))\right\}\mathrm{d}t \\
&\quad + \int \psi\big(\xi_{t-}(x)+\gamma(\xi_{t-}(x),z)\big) - \psi\big(\xi_{t-}(x)\big) - \nabla\psi(\xi_{t-}(x))\gamma'(\xi_{t-}(x))z\ \nu(\mathrm{d}z)\mathrm{d}t \\
&= \nabla\psi(\xi_t(x))\Big(\ \beta(\xi_t) - \int \gamma(\xi_{t-}(x),z) - \gamma'(\xi_{t-}(x))z\ \nu(\mathrm{d}z) \\
&\qquad - \frac{1}{2}\sum a_j^2\nabla\sigma_j(\xi_t(x))\sigma_j(\xi_t(x))\ \Big)\mathrm{d}t \\
&\quad + \frac{1}{2}\Big(\sum \nabla\psi(\xi_t(x))\nabla\sigma_j(\xi_t(x))\sigma_j(\xi_t(x)) + \operatorname{Tr}\left\{\sigma(\xi_t(x))^*\nabla^2\psi(\xi_t(x))\sigma(\xi_t(x))\right\}\Big)\mathrm{d}t \\
&= \frac{1}{2}\sum \sigma_j(\xi_t(x))^*\nabla\left(\nabla\psi(\xi_t(x))\sigma_j(\xi_t(x))\right)\mathrm{d}t = 0\ .
\end{aligned}$$

The last equality follows from $\nabla\psi\sigma_j \equiv 0$ and the previous from (6.9) and the chain rule. ■

For completeness and later reference we also prove the following theorem.

Theorem 6.6. *Under the conditions of the above theorem $\nabla\xi$ leaves $T\mathbb{M}$ invariant. For any $t \geqslant 0$ the restriction of $\nabla\xi_t$ onto $T\mathbb{M}$ homeomorphically maps $T_x\mathbb{M}$ onto $T_{\xi_t(x)}\mathbb{M}$.*

Proof. We want to show that for any $x \in \mathbb{M}, y \in T_x\mathbb{M}, t \geqslant 0$ we have

$$\langle \nabla\psi(\xi_t(x)), \nabla\xi_t(x)y\rangle = 0 \quad \mathbb{P}\text{-a.s.}\ .$$

To this aim consider a smooth curve on a small interval $c : (-\varepsilon,\varepsilon) \to \mathbb{M}$ such that $c(0) = x, \dot{c}(0) = y$. Its image under the flow ξ is the random curve

$$\chi : s \mapsto \chi(s) = \xi_t^x|_{x=c(s)} \in \mathbb{M}\ .$$

Therefore we have $\psi\circ\chi \equiv 0$ almost surely. We may take the derivative at 0 to obtain

$$0 = \frac{\partial}{\partial s}\psi\circ\chi(s)|_{s=0} = \langle \nabla\psi(\xi_t(x)), \nabla\xi_t(x)y\rangle = 0 \quad \mathbb{P}\text{-a.s.}\ .$$

■

Let us apply the above results of this section to the Marcus' canonical equation.

Corollary 6.7 (Marcus' canonical equation)**.** *Consider the Marcus canonical equation* (1.8)

$$\xi_t(x) = x + \sum_j \int_0^t \sigma_j(\xi_{t-}(x)) \diamond \mathrm{d}Z_t^j\ , \tag{6.11}$$

driven by a Lévy process Z with characteristic triplet (b, A, ν) where the vector fields σ_j are all tangent to $\mathbb{M}$, i.e.

$$\sigma_j(x) \in T_x\mathbb{M}, \quad x \in \mathbb{M} \; . \tag{6.12}$$

Then $\xi_t(x) \in \mathbb{M}$ for all $t \geqslant 0$ whenever $x \in \mathbb{M}$, $\mathbb{P}$-almost surely.

Proof. Let us verify the assumptions of the theorem by showing (6.10). By assumption (i) is fulfilled. This also ensures that $x + \gamma(x, z) = \phi^{\sigma z}(x) \in \mathbb{M}$ whenever $x \in \mathbb{M}$, so (ii) is valid. Furthermore by the definition of $\phi^{\sigma z}$ we have $\gamma'(x)z = \nabla_z \gamma(x, z)|_{z=0} z = \sigma(x)z$. Recall that by the Stratonovich correction the drift is given by

$$\beta(x) = \sigma(x)b + \frac{1}{2}\sum \nabla\sigma_j(x)\sigma_j(x) + \int \left(\phi^{\sigma z}(x) - x - \sigma(x)z\right) \; \nu(\mathrm{d}z) \; . \tag{6.13}$$

If we plug β into equation (6.9) reduces to

$$\sigma(x)b = \sum b_j \sigma_{\cdot j} \in T_x\mathbb{M}$$

by assumption. ■

6.2 Smoothness on compact submanifolds

In Chapter 3 we used the integration-by-parts formula to obtain densities and gradient estimates for jump diffusion on Euclidean space $\mathbb{R}^d$. The analysis in this section transfers the results to submanifolds considered in the previous subsection.

While the integration-by-parts formula can be taken from the ambient $\mathbb{R}^d$, gradient estimates and densities are subtle.

Both strategies require the (reduced) Malliavin matrix $\mathscr{A}_t^\theta$ to be invertible. There is no hope to establish invertibility in the ambient $\mathbb{R}^d$ since the dimensions differ. However this is not necessary since we are only allowed to vary the initial condition in directions tangent to the submanifold $\mathbb{M}$.

We consider again the stochastic differential equation (6.1)

$$\begin{cases} \mathrm{d}\xi_t = X(\xi_{t-}, \mathrm{d}t) \; , \\ \xi_0 = x \in \mathbb{M} \; , \end{cases} \tag{6.14}$$

with an $\mathbb{R}^d$ valued semimartingale generator

$$X(x, t) = t\beta(x) + \sigma(x)W_t + \iint_0^t \gamma(x, z)\tilde{N}(\mathrm{d}z\mathrm{d}s) \; . \tag{6.15}$$

We assume that the support of the Lévy measure ν is compact and small enough such that the mappings $x \mapsto x + \gamma(x, z)$ are homeomorphic.

In Section 6.1 we have established conditions such that ξ leaves $\mathbb{M}$ invariant. Thus from now on we assume that the conditions of Theorem 6.4 are met.

Then the flow of diffeomorphisms ξ generated by (6.14) leaves $\mathbb{M}$ invariant and its Jacobian $\nabla\xi$ maps $T\mathbb{M}$ to itself.
As in the case of Euclidean space the core of the argument lies on the relation $\mathscr{D}_\theta\xi_t = \nabla\xi_t\mathscr{A}_t^\theta$ between the Malliavin derivative and the Jacobian of the flow of Theorem 2.14. If $\mathscr{A}_t^\theta$ (*cf.* (5.25)) is invertible, we can transfer the spatial variation of the Jacobian to a variation in the noise by means of the Malliavin derivative and then evoke the integration-by-parts formula.
Let us now restate the ellipticity condition 10 in a version suitable for submanifolds.

Condition 12 (Tangential ellipticity)**.** In the framework of (6.14), assume that the Lévy measure satisfies an order condition with $0 < \alpha < 2$. Denote again

$$\Xi(x) := \sigma^*(x)A\sigma(x) + \gamma'^*(x)A_\nu\gamma'(x) \ . \tag{6.16}$$

We ask that there exists a constant $\kappa_0 > 0$ with

$$\langle v, \Xi(x)v\rangle \geqslant \kappa_0|v|^2 \ , \quad \forall v \in T_x\mathbb{M}, x \in \mathbb{M} \ . \tag{6.17}$$

In extend the result to manifolds we will thoroughly go through the arguments of Chapter 3. We chose again a perturbation θ according to (5.12). On $\mathbb{R}^d$ the invertibility of $\mathscr{A}_t^\theta$ is established in Lemma 5.12. We show that Condition 12 is tailor-made to ensure that lemma and hence the invertibility pertains when restricted to the tangent space.

Lemma 6.8. *Consider the SDE* (6.14) *starting at $x \in \mathbb{M}$ and assume that the conditions of Theorem 6.4 are satisfied. Let further Condition 12 hold. Then*

$$\sup_{v\in T_x\mathbb{M};|v|=1} \mathbb{E}\langle v, \mathscr{A}_t^\theta v\rangle^{-p} < \infty \ , \quad \forall p \geqslant 1, c \in \mathbb{M} \ . \tag{6.18}$$

Proof. The proof goes along the lines of the proof of Lemma 5.12. The pivotal point is Lemma 5.19 and its proof consists of showing (5.38) which is based on the estimate

$$\langle(\nabla\xi_s)^{-1*}v, \Xi(\xi_s)(\nabla\xi_s)^{-1*}v\rangle \geqslant \kappa_0|v^*(\nabla\xi_s)^{-1}|^2 \ .$$

The estimate continues to hold under Condition 12 once we can guarantee that $(\nabla\xi_s)^{-1*}v \in T_{\xi_t}\mathbb{M}$. This however is a consequence of Theorem 6.6.
Indeed, since $\nabla\xi_s$ is invertible and maps $T_x\mathbb{M}$ to $T_{\xi_s}\mathbb{M}$ we have $(\nabla\xi_s)^{-1}y \in T_x\mathbb{M}$ for $y \in T_{\xi_s}\mathbb{M}$. This implies that $(\nabla\xi_s)^{-1}$ maps $T_{\xi_s}\mathbb{M}^\perp$ to $T_x\mathbb{M}^\perp$ and thus

$$\langle\nabla\psi(\xi_s), (\nabla\xi_s)^{-1*}v\rangle = \langle(\nabla\xi_s)^{-1}\nabla\psi(\xi_s), v\rangle = 0 \ .$$

Hence $(\nabla\xi_s)^{-1*}v \in T_{\xi_s}\mathbb{M}$. ■

We continue with the gradient estimate (3.41) thus establishing the strong Feller property of the solution. (The theorem should be compared with Theorem 6.1 in [Mal97, p.245]).

Theorem 6.9. *Assume that Condition 12 holds. Then for every $f \in \mathscr{C}_b^1(\mathbb{M})$, $t > 0$ and $y \in T_x\mathbb{M}, x \in \mathbb{M}$ we have*

$$\nabla\big(\mathcal{P}_t f\big)(x)y = \mathbb{E}\big[\ \nabla f(\xi_t(x))\nabla\xi_t(x)y\ \big] = -\mathbb{E}\big[\ f(\xi_t(x))\Gamma_2\ \big]\ , \tag{6.19}$$

where for all $p > 0$, $\Gamma_2 \in L^p(\Omega)$ is given by $\Gamma_2 = \mathbf{\Gamma}_\theta\left((\mathscr{A}^\theta)^{-1}y\right)$ with $\mathbf{\Gamma}_\theta$ as in (3.35).

Proof. The proof is analogous to the one of Theorem 3.17. Since $\mathscr{A}^\theta$ is invertible on $T_x\mathbb{M}$ we can define $\mathbf{\Psi} = (\mathscr{A}^\theta)^{-1}y$ and use the integration-by-parts formula (3.33). ∎

Corollary 6.10. *Let $\phi : \mathbb{M} \supset U \to V \subset \mathbb{R}^{d-1}$ be a local chart with $x \in U$. Denote by $\tilde{\mathcal{P}}$ the semigroup acting on $\mathscr{C}_c^1(V)$ generated by the process $\zeta_t = \phi(\xi_t)$. Then*

$$\nabla\big(\tilde{\mathcal{P}}_t f\big)(x) = -\mathbb{E}\big[\ f(\zeta_t(x)) \cdot \mathbf{\Gamma}_\theta\Big((\mathscr{A}_t^\theta)^{-1}\nabla\psi(x)v\Big)\ \big]\ , \quad x \in V \tag{6.20}$$

Proof. Denote by ψ the inverse of ϕ. It suffices to note that for any $x, v \in V$ we have that $\nabla\psi(x)v \in T_{\psi(x)}\mathbb{M}$. Then Theorem 6.9 reads

$$\begin{aligned}\nabla\big(\tilde{\mathcal{P}}_t f(x)\big)(v) &= \nabla_v(\mathcal{P}(f\circ\phi)(\psi(x)) = \nabla\mathcal{P}(f\circ\phi)(\psi(x))\nabla\psi(x)v\\ &= \mathbb{E}\big[\ \nabla(f\circ\phi)\big(\xi_t(\psi(x))\big)\cdot\nabla\xi^{\psi(x)}\nabla\psi(x)v\ \big]\\ &= \mathbb{E}\big[\ (f\circ\phi)\big(\xi_t(\psi(x))\big)\cdot\mathbf{\Gamma}_\theta\Big((\mathscr{A}_t^\theta)^{-1}\nabla\psi(x)v\Big)\ \big]\ .\end{aligned}$$

∎

Corollary 6.11. *Under Condition 12 the process ξ restricted to $\mathbb{M}$ is strong Feller.*

Theorem 6.12. *Let $\phi : \mathbb{M} \supset U \to V \subset \mathbb{R}^{d-1}$ be a local chart and let for $x \in \mathbb{M}, t > 0$ the law of $\xi_t(x)$ on $\mathbb{M}$ be denote by μ. Then the image measure $\phi_{\#}\mu = \mu\circ\phi^{-1}$ (the push forward of μ onto V) is absolutely continuous with respect to the $d-1$-dimensional Lebesgue measure on V.*

Proof. We may assume that the Gram determinant $|\det\nabla\phi(x)\nabla\phi(x)^*| \geq c > 0$ on U (otherwise we would shrink U sufficiently). We use the criterion on $\mathbb{R}^{d-1}$. Let $f \in \mathscr{C}_0^\infty(V)$ be given and note that f is compactly supported in V. Denote the pseudo-inverse by

$$R_\phi(x) = \nabla\phi(x)^*(\nabla\phi(x)\nabla\phi(x)^*)^{-1}\ .$$

Then we have by the integration-by-parts formula

$$\begin{aligned}\mu_{\#}\phi(\nabla f) &= \mathbb{E}\big[\ (\nabla f)\circ\phi(\xi_t(x))\ \big] = \mathbb{E}\big[\ \nabla(f\circ\phi)(\xi_t(x))R_\phi(\xi_t(x))\ \big]\\ &= -\mathbb{E}\big[\ (f\circ\phi)(\xi_t(x))\mathbf{\Gamma}_\theta(\mathbf{\Psi})\ \big]\end{aligned}$$

with

$$\mathbf{\Psi} = (\mathscr{D}_\theta\xi_t(x))^{-1}R_\phi(\xi_t(x)) \in \mathbb{R}^{d\times(d-1)}\ . \tag{6.21}$$

∎

Corollary 6.13. *Under Condition 12 the law of ξ_t restricted to $\mathbb{M}$ has a density with respect to the canonical volume on $\mathbb{M}$.*

Part II

Lyapunov exponents for linear equations

Chapter 7

Exponential growth rates. Lyapunov exponents for linear systems

7.1 Lyapunov exponents and dynamical systems

The notion of Lyapunov exponents goes back to his 19th century thesis (see the reprint [Lya92]) and is coined as Lyapunov's first method to describe the stability of dynamical systems. The method describes the exponential growth rate of non-autonomous linear dynamical systems of the form

$$\begin{cases} \dot{y} = B(t)y \ , \\ y(0) = y_0 \in \mathbb{R}^d \ , \end{cases} \tag{7.1}$$

where B is a $d \times d$ matrix valued function defined on $\mathbb{R}_+$. The exponential growth rate of an orbit is then given by the limit (if it exists)

$$\lim_{t\to\infty} \frac{1}{t} \ln |y(t)| \ . \tag{7.2}$$

More generally, the maps $B(t)$ may form a *cocycle of linear maps* on some normed vector space $\mathbb{X}$ over a base dynamical system $(\vartheta_t)_{t\in\mathbb{T}}$ on some topological space Ω. Let us specify this definition (for the sake of this outline we set aside questions of measurability).
We may investigate the dynamics with respect to discrete or continuous, one-sided or two-sided notions of time symbolized by an index set $\mathbb{T}$. Hence $\mathbb{T}$ may be taken to be equal to either of the ordered sets $\mathbb{N}_0, \mathbb{Z}, \mathbb{R}, \mathbb{R}_+$. A family of transformations $\vartheta = (\vartheta_t)_{t\in\mathbb{T}}$ of the space Ω is said to be a *dynamical system* if it satisfies

- $\vartheta_0 = \mathrm{id}_\Omega$ is the identity on Ω,
- the *(semi-) flow* property: $\vartheta_{t+s} = \vartheta_t \circ \vartheta_s, \ s,t \in \mathbb{T}$.

In addition we may say that the dynamical system ϑ is *metric* if there exists an invariant measure μ on Ω for the transformations $\vartheta_t, t \in \mathbb{T}$. Equivalently ϑ is *measure preserving* with respect to μ, i.e.

$$(\vartheta_t)_\# \mu(A) = \mu(\vartheta_t^{-1}(A)) = \mu(A)$$

for any $\mu-$measurable subset A of Ω and $t \in \mathbb{T}$ (obviously this requires measurability of ϑ). A map $\varphi : \mathbb{T} \times \Omega \to L(\mathbb{X})$ into the linear transformations of $\mathbb{X}$ then forms a *linear cocycle* over ϑ if it satisfies

- $\varphi(0, \omega) = \mathrm{id}_{\mathbb{X}}$ is the identity on $\mathbb{X}$,
- the *cocycle* property: $\varphi(t + s, \omega) = \varphi(t, \vartheta_s(\omega)) \circ \varphi(s, \omega),\ s, t \in \mathbb{T}$.

Predominant examples of such linear cocycles (see [BDV05, Via14]) are

1. *Jacobian cocycles*. Consider an *ODE* on $\mathbb{R}^d$ of the form $\dot{x} = X(x)$ for a smooth and complete vector field X on $\mathbb{R}^d$. The generated flow ϑ is a dynamical system on $\Omega = \mathbb{R}^d$. The Jacobian matrix $\nabla\vartheta$ of ϑ with

$$(\nabla\vartheta(x))^{ij} = \frac{\partial\vartheta^i}{\partial x^j}(x)$$

   then forms a linear cocycle over ϑ with $\mathbb{T} = \mathbb{R}$. Note that if there exists an invariant measure μ for the dynamics of ϑ it is metric – although μ is in general not finite. To obtain finite measures one should consider a compact state space Ω, e.g. a compact submanifold of $\mathbb{R}^d$.

2. *Random matrices*. Let $B_1, B_2, \cdots$ be a sequence of *i.i.d.* random matrices. If a random matrix B_0 on a probability space $(\Omega_0, \mathbb{P}_0)$ has the same distribution as B_1 we may model the sequence on the canonical probability space $\Omega = \Omega_0^{\mathbb{N}}$ with the product measure $\mathbb{P} = \mathbb{P}_0^{\otimes\mathbb{N}}$. With $\mathbb{T} = \mathbb{N}$ the base dynamical system is given by the canonical shift

$$\begin{aligned} \vartheta : \Omega &\longrightarrow \Omega \\ (\omega_1, \omega_2, \ldots) &\mapsto (\omega_2, \omega_3, \ldots) \end{aligned} \tag{7.3}$$

   and the iteration $\vartheta_n(\omega) = \vartheta^n(\omega) = \vartheta \circ \cdots \circ \vartheta(\omega)$ (n-times). Note that ϑ is $\mathbb{P}$-measure preserving and thus metric. On this probability space we have that $B_n = B_0(\vartheta_n(\cdot))$. The product of the random matrices $A_n = B_1 \cdot B_2 \cdots B_n$ then forms a linear cocycle over ϑ.

3. *Jacobians of random transformations.* This case is a combination of the two preceding examples. Assume that on the probability space $(\Omega_0, \mathbb{P}_0)$ we have a random transformation $f : \mathbb{M} \to \mathbb{M}$, where for simplicity $\mathbb{M}$ is a compact submanifold of $\mathbb{R}^d$. Assume further that the random transformation f has an invariant measure μ on $\mathbb{M}$ under $\mathbb{P}_0$ and that f is almost surely differentiable on $\mathbb{M}$. On the canonical product space $(\Omega, \mathbb{P})$ as above we can define the random orbit $f_n(\omega, x) = f(\omega_n, f(\omega_{n-1}, \cdots f(\omega_1, x)) \cdots), x \in \mathbb{M}$. It is easy to see that the random orbit $(f_n)_{n\in\mathbb{N}}$ is a (non-linear) cocycle over the (metric) dynamical system ϑ of (7.3).

   Similarly we can define the Jacobian of f_n which due to the chain rule satisfies

$$\nabla f_n(\omega, x) = \nabla f(\omega_n, f_{n-1}(\omega, x))\nabla f_{n-1}(\omega, x)\ . \tag{7.4}$$

It follows that the collection $\big(\nabla f_n\big)_{n\in\mathbb{N}}(\omega,x)$ forms a linear cocycle over the *skew-product flow*

$$\Theta_n(\omega,x) = (\omega_n, f_n(\omega,x)) \text{ on } \Omega\times\mathbb{M}\ . \tag{7.5}$$

Also note that by construction the dynamical system $\Theta = (\Theta_n)_{n\in\mathbb{N}}$ is measure preserving with respect to the product measure $\mathbb{P}\otimes\mu$ on $\Omega\times\mathbb{M}$ and hence is metric.

For the remainder of this thesis we investigate linear cocycles of the form 2. where the random matrix structure is derived from the stochastic flow generated by linear *SDE* driven by Lévy processes on Wiener–Poisson space and in the Euclidean space $\mathbb{R}^d$.

7.2 Lyapunov exponents of linear *SDE*

We investigate the exponential growth rate

$$\lambda = \lambda(x) = \lim_{t\to\infty}\frac{1}{t}\ln|\xi_t(x)|\ ,\quad x\in\mathbb{R}^d\ , \tag{7.6}$$

of a linear version of the stochastic differential equation (1.14). More precisely we consider a linear *SDE* of the form

$$\begin{cases} \mathrm{d}\xi_t = B_0\xi_t\mathrm{d}t + \displaystyle\sum_{j=1}^m B_j\xi_{t-}(\diamond)\mathrm{d}Z_t^j\ , \\ \xi_0 = x\in\mathbb{R}^d\ , \end{cases} \tag{7.7}$$

where the stochastic integral is taken to be either of multiplicative or of Marcus' canonical type (see Section 1.2.2). Here B_0, B_j $(j=1,\dots,m)$ are deterministic matrices in $\mathbb{R}^{d\times d}$ and the driving Lévy process Z with values in $\mathbb{R}^m$ has the characteristic triplet denoted by (b,A,ν). Without loss of generality we may assume that $A = \operatorname{diag}(a_1^2, a_2^2,\dots,a_m^2)$ with $a_j \geqslant 0$. Note that the coefficients are trivially of linear growth and Lipschitz, such that there exists a unique strong solution (*cf.* Condition 1). Our main focus lies on the derivation of a Furstenberg–Khasminskii type representation of (7.6) in Chapter 8.

Remark 7.1 (Nonlinear equations)*.* In general the coefficient matrices could be previsible matrix processes $B_0(t), B_j(t)$ $(j=1,\dots,m)$. In the stability analysis of non-linear equations one studies the growth of the linearization of the flow. Indeed, we have shown that the linearization (the Jacobian) follows (2.15) which is a non-autonomous matrix version of (7.7). Stability is then naturally associated with the negativity of (7.6). Apart from some remarks we will only consider linear equations in this thesis.

Remark 7.2*.* Lyapunov exponents for linear *SDE* with jumps have been studied by various authors [BL86a, BL86b], [Mao99]. Linear jump equations play also an important role in the context of Markov switching [Mao99, KZY07, YY12, LMR14, CH15] and the monograph [CFT13]. [AS09, AS10] consider exponential growth rates in the light of stabilization avoiding the concept of linearization in the spirit of [Mao94].

Lyapunov exponents for functional equations driven by general semimartingales have been studied in [MS96, MS97].
[LDS02] derive a Furstenberg–Khasminskii formula for equations of type (7.7) if the jumps have a first moment and thus avoiding the Itô integral. Their method relies on the smoothing properties of a Gaussian component. A Furstenberg–Khasminskii formula in a similar context is also found in [Sko89]. A very comprehensive geometric treatment can be found in [Lia04].

Let us specify the meaning of the two interpretations.

Multiplicative type. If we consider the stochastic integral of multiplicative type equation (7.7) can be expressed in the semimartingale-generator form

$$\begin{cases} \mathrm{d}\xi_t = X(\xi_{t-}, \mathrm{d}t) = \boldsymbol{X}(\mathrm{d}t)\xi_{t-} \ , \\ \ \xi_0 = x \in \mathbb{R}^d \ . \end{cases} \tag{7.8}$$

By linearity $X(x,t) = \boldsymbol{X}(t)x$ with a matrix valued semimartingale generator $\boldsymbol{X}(t)$ represented by

$$\begin{aligned} X(t,x) &= t(B_0 + \sum_{j=1}^m b^j B_j)x + \sum_{j=1}^m a_j W_t^j B_j x + \iint_{\mathbb{B}\times[0,t]} \sum_{j=1}^m z^j B_j x \tilde{N}(\mathrm{d}z\mathrm{d}s) \\ &\quad + \iint_{\mathbb{B}^c\times[0,t]} \sum_{j=1}^m z^j B_j x N(\mathrm{d}z\mathrm{d}s) \\ &= (tB_0 + \sum_{j=1}^m Z_t^j B_j)x = \boldsymbol{X}(t)x \ , \qquad x \in \mathbb{R}^d, t \geqslant 0 \ . \end{aligned}$$

In fact $\boldsymbol{X}(t)$ is a matrix valued Lévy process. The linear integrand is actually a coordinate representation of a bilinear map (a (1-1)-tensor) $\boldsymbol{B} : \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}^d$. We will often rely on this representation for better readability. We denote

$$z^* \boldsymbol{B} x = \sum_{j=1}^m z^j B_j x \ , \quad z \in \mathbb{R}^m, x \in \mathbb{R}^d. \tag{7.9}$$

The reader may think of $\boldsymbol{B}$ as a column vector with matrix entries B_j. With this notation we have

$$\boldsymbol{X}(t) = tB_0 + tb^*\boldsymbol{B} + (A^{\frac{1}{2}}W_t)^*\boldsymbol{B} + \iint_{\mathbb{B}\times[0,t]} z^*\boldsymbol{B}\tilde{N}(\mathrm{d}z\mathrm{d}s) + \iint_{\mathbb{B}^c\times[0,t]} z^*\boldsymbol{B}N(\mathrm{d}z\mathrm{d}s) \ , \tag{7.10}$$

$$= tB_0 + Z_t^*\boldsymbol{B} \ . \tag{7.11}$$

We will also state the infinitesimal generator $\mathcal{L}$ of the Markov semigroup associated to (7.7). For any $f \in \mathscr{C}_c^\infty(\mathbb{R}^d)$ we have

$$\begin{aligned}\mathcal{L}f(x) = \langle B_0 x, \nabla\rangle f(x) + \langle b^* \boldsymbol{B} x, \nabla\rangle f(x) + \frac{1}{2} a_j^2 (\langle B_j x, \nabla\rangle f)^2(x) \\ + \int f((\mathrm{Id} + z^* \boldsymbol{B})x) - f(x) - \mathbb{1}_{|z|\leqslant 1} \nabla f(x)(z^* \boldsymbol{B} x)\nu(\mathrm{d}z) \ .\end{aligned} \tag{7.12}$$

Here we used Einstein's summation convention. In particular this notation means

$$a_j^2(\langle B_j x, \nabla\rangle f)^2(x) = \sum_j \sum_{khmn} a_j^2 (B_j)_{kh} (B_j)_{mn} x_k x_m \partial_h \partial_n f(x), \ x \in \mathbb{R}^d \ . \tag{7.13}$$

So ∇ acts only on f and the square is a regular square on real numbers.

Marcus' canonical type. The concept of *SDE*of canonical type has been discussed in Section 1.2.2. If we interpret (7.7) in the canonical sense it can be represented as

$$\xi_t(x) = x + \int_0^t X^\diamond(\xi_{s-}(x), \mathrm{d}s) \ . \tag{7.14}$$

The semimartingale generator is given by

$$\begin{aligned}X^\diamond(t,x) = t\big\{B_0 + \sum_{j=1}^m \big(b^j B_j + \frac{1}{2} a_j^2 B_j B_j\big) + \int_{\mathbb{B}} (\phi^{z^*\boldsymbol{B}} - \mathrm{Id} - \sum_{j=1}^m z^j B_j) \ \nu(\mathrm{d}z)\big\} x \\ + \sum_{j=1}^m \int_0^t a_j W_t^j B_j x + \iint_{\mathbb{B}\times[0,t]} (\phi^{z^*\boldsymbol{B}} - \mathrm{Id}) x \ \tilde{N}(\mathrm{d}z\mathrm{d}s) \\ + \iint_{\mathbb{B}^c\times[0,t]} (\phi^{z^*\boldsymbol{B}} - \mathrm{Id}) x \ N(\mathrm{d}z\mathrm{d}s) \ .\end{aligned} \tag{7.15}$$

In view of (1.29) the equation is interpreted as follows. The solution jumps along the deterministic curve $\phi : [0,1] \to \mathbb{R}^d$ solving

$$\begin{cases} \dot{\phi} = \sum_{j=1}^m z^j B_j \phi = z^* \boldsymbol{B} \phi \ , \\ \phi_0 = x \ . \end{cases} \tag{7.16}$$

In particular we denote by $\phi^{z^*\boldsymbol{B}}$ the matrix representation of the time-one mapping $x \mapsto \phi_1 x$. By linearity we have again the representation by a matrix valued semimartingale $X^\diamond(t,x) = \boldsymbol{X}^\diamond(t)x$, this time non-linear in the noise. Indeed, we have

$$\boldsymbol{X}^\diamond(t) = t\big(B_0 + b^* \boldsymbol{B} + \frac{1}{2} \mathbf{Tr}\,|_{\mathbb{R}^{d\times d}} \boldsymbol{B}^* A \boldsymbol{B} + \int_{\mathbb{B}} (\phi^{z^*\boldsymbol{B}} - \mathrm{Id} - z^* \boldsymbol{B}) \ \nu(\mathrm{d}z)\big) \tag{7.17}$$

$$+ (A^{\frac{1}{2}} W_t)^* \boldsymbol{B} + \iint_{\mathbb{B}\times[0,t]} (\phi^{z^*\boldsymbol{B}} - \mathrm{Id}) \ \tilde{N}(\mathrm{d}z\mathrm{d}s) \tag{7.18}$$

$$+ \iint_{\mathbb{B}^c\times[0,t]} (\phi^{z^*\boldsymbol{B}} - \mathrm{Id}) \ N(\mathrm{d}z\mathrm{d}s) \ , \ t \geqslant 0 \ . \tag{7.19}$$

where $\mathbf{Tr}\,|_{\mathbb{R}^{d\times d}}$ is the tensor contraction (trace) to $d\times d$-matrices ([MRA07, p.344] or [BG80, pp.85-86]).

Remark 7.3. The driving matrix-valued semimartingale generator of the two equations correspond to each other via the Itô–Stratonovich conversion

$$\boldsymbol{X}^{\diamond}(t) = \boldsymbol{X}(t) + t\frac{1}{2}\,\mathbf{Tr}\,|_{\mathbb{R}^{d\times d}}\boldsymbol{B}^{*}A\boldsymbol{B} + \iint_{0}^{t} (\phi^{z^{*}\boldsymbol{B}} - \mathrm{Id} - z^{*}\boldsymbol{B})\; N(\mathrm{d}z\mathrm{d}s)\;, \tag{7.20}$$

where $\boldsymbol{X}$ is the matrix valued semimartingale defined in (7.10).

Lemma 7.4. *Assume that the matrices $B_j, j = 0, \ldots, m$, commute. Then the solution flow ξ of* (7.14) *is given by the matrix exponential*

$$\xi_t(x) = \exp\big(tB_0 + Z_t^{*}\boldsymbol{B}\big)x\;. \tag{7.21}$$

Proof. Observe that if the matrices commute, then the fundamental solution to (7.16) is given by the matrix exponential

$$\exp\big(tz^{*}\boldsymbol{B}\big)\;. \tag{7.22}$$

Therefore the jump kernel of (7.15) equals $(\exp\big(tz^{*}\boldsymbol{B}\big) - \mathrm{Id})x$. The Gaussian part consists of a Stratonovich integral and we may rely on a Wong–Zakai approximation argument. Indeed, if $t_n, n = 0, \ldots, N$, is a partition of $[0, T], T > 0$ of mesh size $\max_n \delta t_n \leqslant \delta$ for some $\delta > 0$ we denote by $\delta_n W(t) = W_{t_n} - W_{t_{n-1}}$ if $t_{n-1} < t \leqslant t_n$. Let ξ_t^{δ} be the solution to the piece wise constant linear random *ODE*

$$\dot{\phi} = \delta_n W^{*}\boldsymbol{B}\phi\;. \tag{7.23}$$

Since the (piece wise) fundamental solution is given by (7.22) with $z = \delta_n W$ we deduce that

$$\xi_T^{\delta} = \prod_n \exp\big(\delta_n W^{*}\boldsymbol{B}\big) = \exp\Big(\sum_n \delta_n W^{*}\boldsymbol{B}\Big) = \exp\big(W_T^{*}\boldsymbol{B}\big)\;. \tag{7.24}$$

■

We deduce that the infinitesimal generator $\mathcal{L}^{\diamond}$ of the Markov semigroup associated to canonical equation (7.8) is given by

$$\begin{aligned}\mathcal{L}^{\diamond} f(x) &= \mathcal{L}f(x) + \frac{1}{2}a_j^2\langle B_jB_jx, \nabla\rangle f(x) + \int f(\phi^{z^{*}\boldsymbol{B}}(x)) - f(x) - \nabla f(x)(z^{*}\boldsymbol{B}x)\nu(\mathrm{d}z)\\ &= \langle B_0x, \nabla\rangle f(x) + \frac{1}{2}a_j^2(\langle B_jx, \nabla\rangle)^2 f(x)\\ &\quad + \int f(\phi^{z^{*}\boldsymbol{B}}(x)) - f(x) - \mathbb{1}_{|z|\leqslant 1}\nabla f(x)(\phi^{z^{*}\boldsymbol{B}}(x) - x)\nu(\mathrm{d}z)\;.\end{aligned} \tag{7.25}$$

Note that the local part of the operator is of Hörmander form. Indeed, if we denote the linear vector fields by $v_j(x) = a_jB_jx, j = 1, \ldots, m$, then the local part of (7.25) reads

$$\sum_j v_j^2(x)\;,$$

where the square is the square (iteration) of operators.

7.3 The Multiplicative Ergodic Theorem (MET)

The exponential growth rate of a linear *SDE* is closely related to Oseledec's celebrated *Multiplicative ergodic theorem* (MET) [Ose68]. Indeed (7.6) is the asymptotic limit of the maximal logarithmic eigenvalue of the linear operator given by the stochastic flow

$$x \mapsto \xi_t(x) \ .$$

The MET provides a spectral theory for linear cocycles provided Ruelle's integrability condition for the positive part of the logarithm $\mathbb{E}\big[\ \ln^+ |\xi(x)|\ \big] < \infty$ holds (where $\ln^+ = \max\{0, \ln(\cdot)\}$ is the positive part of the logarithm). The condition rules out infinity as an eigenvalue of the flow which would indicate superexponential growth.
We state the MET for discrete (one-sided) time following [GM89, Thm.1.2] (*cf.* [Via14, Thm.4.1], also [Arn98, Thm.3.4.1]).

Theorem 7.5 (Oseledec's MET)**.** *Let $\vartheta : \Omega \to \Omega$ be a $\mathbb{P}$-ergodic shift on Ω, $\Phi : \Omega \to GL(\mathbb{R}^d)$ be a measurable map into the group of invertible $d \times d$ matrices. Suppose that*

$$\ln^+ \|\Phi\| \in L^1(\Omega, \mathbb{P}) \ . \tag{7.26}$$

We define the matrix cocycle $\Phi_n(\omega) := \Phi(\vartheta^{n-1}\omega) \cdot \Phi(\vartheta^{n-2}\omega) \cdots \Phi(\omega)$. Then there exists a forward invariant set $\tilde{\Omega} \subset \Omega$ of full measure such that

(i) If $\eta(n,\omega) \leqslant \cdots \leqslant \eta(n,\omega)$ are the eigenvalues of the random matrix $(\Phi_n^*\Phi_n)^{\frac{1}{2n}}(\omega)$ then the limits

$$\lim_{n\to\infty} \frac{1}{n} \log \eta_i(n,\omega) =: \tilde{\lambda}_i(\omega) \ , \quad i = 1, \dots, d \ , \tag{7.27}$$

exist with probability 1.

(ii) The limit

$$\lim_{n\to\infty} (\Phi_n^*\Phi_n)^{\frac{1}{2n}}(\omega) =: \Lambda(\omega) \ , \quad i = 1, \dots, d \ , \tag{7.28}$$

exist with probability 1.

(iii) The set numbers $\mathrm{e}^{\tilde{\lambda}_i}(\omega), i = 1, \dots, d$ coincide with the eigenvalues of $\Lambda(\omega)$.

(iv) Denote by $(\lambda_i, k_i), i = 1, \dots, p$, for some $1 \leqslant p \leqslant d$ the set of distinct values of $\tilde{\lambda}_i$ together with their multiplicities. Then the set $((\lambda_1, k_1), (\lambda_2, k_2), \dots, (\lambda_p, k_p))$ called the LYAPUNOV SPECTRUM *is constant a.s..*

(v) Denote by $\mathscr{U}_i(\omega), i = 1, \dots, p$, the eigenspaces corresponding to the eigenvalues e^{λ_i} of $\Lambda(\omega)$. Clearly $\dim \mathscr{U}_i = k_i$. We define the (forward) FLAG

$$0 =: \mathscr{V}_0 \subset \mathscr{V}_1 \subset \cdots \subset \mathscr{V}_p = \mathbb{R}^d, \ \text{with } \mathscr{V}_i = \bigoplus_{j \leqslant i} \mathscr{U}_j \ . \tag{7.29}$$

Then for any $x \in \mathscr{V}_i(\omega) \setminus \mathscr{V}_{i-1}(\omega)$ with probability 1

$$\lim_{n\to\infty} \frac{1}{n} \ln \|\Phi_n(\omega)x| = \lambda_i \ . \tag{7.30}$$

This section verifies the validity of Ruelle's condition (7.26) for the discrete time stochastic flow generated by (7.7) and hence the MET.

7.3.1 Integrability of the multiplicative equation

Here the jumps of the driving Lévy process act multiplicatively on the state of the equation. To ensure Ruelle's integrability condition it is therefore natural to require the finiteness of the logarithmic moment of the Lévy measure. Indeed, we have

Proposition 7.6. *Ruelle's integrability condition* $\mathbb{E}\big[\ \ln^+|\xi_t(x)|\ \big] < \infty$ *holds for all* $t > 0$ *and every initial condition* $x \in \mathbb{R}^d$ *if the Lévy measure satisfies*

$$\int \ln^+|z|\ \nu(\mathrm{d}z) < \infty\ . \tag{7.31}$$

To see that (7.31) is a sharp condition consider the equation with $m = d = 1, B_1 = 1 \in \mathbb{R}^{1\times 1}$ driven by a one dimensional compound Poisson process Z with unit intensity and a Lévy measure ν concentrated on $\mathbb{R}_+$. Then the solution to (7.8) is given by

$$\xi_t(x) = x\prod_{s\leqslant t}(1+\Delta Z_s)\ . \tag{7.32}$$

Consequently Ruelle's condition for $x > 1$ reads

$$\mathbb{E}\ln^+\xi_t(x) = \ln x + t\int \ln(1+z)\nu(\mathrm{d}z) \geqslant t\int \ln^+|z|\nu(\mathrm{d}z)\ . \tag{7.33}$$

It remains to show that the condition is sufficient.

Proof. For the sake of notation assume that $|\xi_0| = 1$. Let $\chi : \mathbb{R}^+ \to [0,1]$ be a smooth monotone function such that $\chi(x) \equiv 0, 0 < x < \mathrm{e}^{-\frac{1}{2}}$ and $\chi(x) \equiv 1, x \geqslant 1$. Then $\ln^+(x) \leqslant |\chi(x)\ln(x)| \leqslant \frac{1}{2} \vee \ln^+(x)$. We observe that the gradient satisfies

$$\frac{1}{2}\big(\chi \times \ln|\cdot|^2\big)'(x) = \frac{1}{2}\chi'(x)\ln(|x|^2) + \chi(x)\frac{x^*}{|x|^2}\ . \tag{7.34}$$

Since all derivatives of χ have bounded support not containing 0 it follows that there is a function $r > 0$ of bounded support such that

$$|\left(\chi\times\ln\right)'(x) - \chi(x)\ln'(x)| + |\left(\chi\times\ln\right)''(x) - \chi(x)\ln''(x)| \leqslant r(x)\ . \tag{7.35}$$

Also

$$\chi(x)\times\ln(x) - \chi(y)\times\ln(y) \leqslant \tfrac{1}{2}\vee\ln^+(x/y)\ , \quad x,y > 0\ . \tag{7.36}$$

To see the last inequality note that if $x, y \geqslant 1$ or $x \geqslant y > 0$, then the left hand side is dominated by $\ln(x/y)$. If $y \geqslant 1 \geqslant x > 0$ it is negative and for $1 \geqslant y \geqslant x > 0$ we obtain

$$\chi(x)\times\ln(x) - \chi(y)\times\ln(y) \leqslant \chi(y)\ln(1/y) \leqslant 1/2\ . \tag{7.37}$$

We apply Itô's formula to $\frac{1}{2}\big(\chi \times \ln\big)(|\xi_t|^2)$. By the observation (7.35) the relevant contributions to Itô's formula derive from the derivatives of $\ln|\cdot|^2$. It follows that there are bounded functions $R_i : \mathbb{R}^d \to \mathbb{R}$ in which all terms containing derivatives of χ are encoded such that by Itô's formula

$$\frac{1}{2}\big(\chi \times \ln\big)(|\xi_t|^2) = \int_0^t \big(R_1(\xi_t) + \chi(|\xi_t|^2)\langle\bar{\xi}_s, B_0\bar{\xi}_s + b^*\boldsymbol{B}\bar{\xi}_s\rangle\big)\mathrm{d}s \tag{7.38}$$

$$+ a_j \int_0^t \big(R_2(\xi_t) + \chi(|\xi_t|^2)\langle\bar{\xi}_s, B_j\bar{\xi}_s\rangle\big)\mathrm{d}W_s^j \tag{7.39}$$

$$+ \int_0^t a_j\chi(|\xi_t|^2)\left(|B_j\bar{\xi}_s|^2 - 2\langle\bar{\xi}_s, B_j\bar{\xi}_s\rangle^2\right)\mathrm{d}s \tag{7.40}$$

$$+ \iint_{\mathbb{B}^c\times[0,t]} \chi \times \ln(|(\mathrm{Id} + z^*\boldsymbol{B})\xi_{s-}|^2) - \chi \times \ln(|\xi_{s-}|^2)N(\mathrm{d}z\mathrm{d}s) \tag{7.41}$$

$$+ \iint_{\mathbb{B}\times[0,t]} \chi \times \ln(|(\mathrm{Id} + z^*\boldsymbol{B})\xi_{s-}|^2) - \chi \times \ln(|\xi_{s-}|^2)\tilde{N}(\mathrm{d}z\mathrm{d}s) \tag{7.42}$$

$$+ \iint_{\mathbb{B}^c\times[0,t]} \Big\{\chi \times \ln(|(\mathrm{Id} + z^*\boldsymbol{B})\xi_{s-}|^2) - \chi \times \ln(|\xi_{s-}|^2) \tag{7.43}$$

$$- \langle\nabla(\chi \times \ln(|\cdot|^2))(\xi_{s-}), z^*\boldsymbol{B}\xi_{s-}\rangle\Big\}\ \nu(\mathrm{d}z)\mathrm{d}s. \tag{7.44}$$

The integrand under the compensator integral with argument $x = \xi_{s-}$ is of the form

$$\begin{aligned}
&\chi \times \ln(|(\mathrm{Id} + z^*\boldsymbol{B})x|^2) - \chi \times \ln(|x|^2) \\
&\quad - \big((\chi \times \ln)'(|x|^2)\big)\left(|(\mathrm{Id} + z^*\boldsymbol{B})x|^2 - |x|^2\right) \\
&\quad + \big((\chi \times \ln)'(|x|^2)\big)\left(|(\mathrm{Id} + z^*\boldsymbol{B})x|^2 - |x|^2 - 2\langle x, z^*\boldsymbol{B}x\rangle\right)\ ,
\end{aligned}$$

where we have regrouped the terms in such a way that the first three terms are a Taylor expansion of $\chi \times \ln$ at $|x|^2$ and the last term contains an expansion of $|\cdot|^2$ at x. We apply the intermediate value theorem twice to see that there is a number h between $|x|$ and $|(\mathrm{Id} + z^*\boldsymbol{B})x|$ and $0 \leqslant \varepsilon \leqslant 1$ such that the sum is equal to

$$\frac{1}{2}\left((\chi \times \ln)''(h^2)\right)(2\langle\varepsilon(\mathrm{Id} + z^*\boldsymbol{B})x + (1-\varepsilon)x, z^*\boldsymbol{B}x\rangle)^2 \tag{7.45}$$

$$+ \left((\chi \times \ln)'(|x|^2)\right)|z^*\boldsymbol{B}x|^2\ . \tag{7.46}$$

Noting that the function r in (7.35) has bounded support, the absolute value of the integrand is bounded up to a multiplicative constant by

$$\chi(h^2)2\langle(\mathrm{Id} + \varepsilon z^*\boldsymbol{B})x/h, z^*\boldsymbol{B}x/h\rangle^2 + \chi(|x|^2)|z^*\boldsymbol{B}\bar{x}|^2\ .$$

This is of order $O(|z|^2)$ uniformly in x. Hence the compensator integral is bounded uniformly in $x = \xi_{s-}$. Under expectation the two martingale terms vanish. The two remaining

$\mathrm{d}t$ terms are integrals of bounded functions. Thus, there is now another constant, again denoted by C, such that

$$\mathbb{E}[\ \chi \times \ln(|\xi_t|^2)] \leqslant Ct + \mathbb{E} \iint_{\mathbb{B}^c \times [0,t]} \tfrac{1}{2} \vee \ln^+(|(\mathrm{Id} + z^* \boldsymbol{B})\bar{\xi}_{s-}|^2) N(\mathrm{d}z\mathrm{d}s) \tag{7.47}$$

$$\leqslant Ct + \tfrac{1}{2} + \mathbb{E} \iint_{\mathbb{B}^c \times [0,t]} 2 \ln^+(1 + \sum |z^j| \|B_j\|) N(\mathrm{d}z\mathrm{d}s) \ , \tag{7.48}$$

which is clearly finite under $\int \ln^+ |z| \mathrm{d}\nu < \infty$. ■

Remark 7.7. In general a Lévy process Z with characteristic (b, A, ν) has a logarithmic moment $\mathbb{E} \ln^+ |Z_t| < \infty$ if and only if ν has a logarithmic moment $\int \ln^+ |z| \ \nu(\mathrm{d}z) < \infty$ (see [Sat99, Thm 25.3]). We can interpret the above as follows. Recall that $\boldsymbol{X}(t)$ is a $\mathbb{R}^{d\times d}$ matrix valued Lévy process with Lévy measure given by the image of ν under the jump kernel. Finiteness of the logarithmic moment with respect to the Frobenius norm, which is the Euclidean vector norm on $\mathbb{R}^{d^2}$, then follows from the finiteness of the logarithmic moment of the image Lévy measure.
Since the induced operator norm of $\boldsymbol{X}(t)$ is bounded by the Frobenius norm (via Cauchy-Schwartz) we also deduce a logarithmic moment of the operator norm. This is however only a formal indicator since the matrix valued Lévy process $\boldsymbol{X}(t)$ does not coincide with the flow generated by (7.7).

7.3.2 Integrability of the canonical equation

In the canonical interpretation the jumps of the Poisson random measure act exponentially on the state. It turns out that the adequate condition to ensure Ruelle's integrability condition is to ask for the finiteness of the first moment of the Lévy measure away from zero.

Proposition 7.8. *Ruelle's integrability condition* $\mathbb{E}\big[\ \ln^+ |\xi_t(x)|\ \big] < \infty$ *is fulfilled for all* $t > 0$ *and any initial condition* $x \in \mathbb{R}^d$ *if the Lévy measure satisfies*

$$\int (|z| - 1)_+ \ \nu(\mathrm{d}z) < \infty \ . \tag{7.49}$$

Remark 7.9. The condition (7.49) is trivially equivalent to $\int_{\mathbb{B}^c} |z| \nu(\mathrm{d}z) < \infty$ for some (and hence any) ball around 0.
We argue that (7.49) is sharp for the equation with $B_1 = 1 \in \mathbb{R}^{1\times 1}$ driven by a one dimensional compound Poisson process Z with unit intensity and a Lévy measure ν concentrated on $\mathbb{R}_+$. Then the solution to (7.8) is given by

$$\xi_t(x) = x\,\mathrm{e}^{Z_t} \ . \tag{7.50}$$

Consequently Ruelle's condition for $x > 1$ reads

$$\mathbb{E}\ln^+ \xi_t(x) = \ln x + t \int z\nu(\mathrm{d}z) \geqslant t \int (|z| - 1)_+ \nu(\mathrm{d}z) \ . \tag{7.51}$$

It remains to show that the condition is sufficient.

Proof. Similarly to the proof of Proposition 7.6 we can use the smoothed version $\chi \times \ln|\cdot|^2$. By Leibniz' rule the calculations are easier in the Marcus case. Only the expectation of the large jump part remains. Recall that $\|\phi^{z^*\boldsymbol{B}}\| \leqslant \exp(\sum |z^j| \|B_j\|)$. We have

$$\mathbb{E}[\ \chi \times \log(|\xi_t|^2)] \leqslant Ct + \mathbb{E} \iint_{\mathbb{B}^c \times [0,t]} \tfrac{1}{2} \vee \ln^+(|\phi^{z^*\boldsymbol{B}} \bar{\xi}_{s-}|^2) N(\mathrm{d}z\mathrm{d}s) \tag{7.52}$$

$$\leqslant Ct + \tfrac{1}{2} + 2\mathbb{E} \iint_{\mathbb{B}^c \times [0,t]} \sum_j |z^j| \|B_j\|) N(\mathrm{d}z\mathrm{d}s)\ , \tag{7.53}$$

which is finite if $\int_{\mathbb{B}^c} |z| \mathrm{d}\nu < \infty$. This condition is an equivalent formulation of (7.49). ■

Chapter 8

A Formula of Furstenberg–Khasminskii type

It follows from the linearity of the flow and additivity of the logarithm that the limiting growth rate (7.6) is invariant under scaling $x \mapsto \alpha x, \alpha \in \mathbb{R} \setminus \{0\}$. Indeed, in view of the MET, it depends only on the Oseledets' (linear) subspace. It is therefore natural to expect that the exponential growth rate is given as an ergodic average of the infinitesimal growth over all directions. This reasoning has been investigated for processes on (Lie-) groups by *Furstenberg* [Fur63] and independently for stochastic differential equations by *Khasminskii* [Kha67] well before Oseledets' theorem was proven. Indeed, we will show in Theorem 8.13 below that

$$\boldsymbol{\lambda} = \int_{\mathbb{S}^{d-1}} \frac{1}{2} \mathcal{L}^{(\diamond)} \big(\ln |\cdot|^2 \big)(\bar{x}) \mu(\mathrm{d}\bar{x}) \ , \tag{8.1}$$

where $\mathcal{L}^{(\diamond)}$ is the generator of the Markov semigroup associated to (7.7) and μ the invariant measure of its projection on the sphere. We refer to this expression as the *Furstenberg–Khasminskii* formula.

In this Chapter we give a proof based on the ergodic theorem of Birkhoff. In fact, since $\boldsymbol{\lambda}$ may be $-\infty$ we will utilize an approximation based on Kingman's subadditive ergodic theorem which to the best of the author's knowledge is new to the literature. We then apply the results of Part I to ensure the uniqueness of the invariant measure μ. A similar formula for the case that Z is a Brownian motion with an additional compound Poisson process (of finite expectation) has been derived in [LDS02].

Let us make some preliminary definitions. We denote by $h_j(x)$ the image of the linear vector field B_j under the projection to the sphere $\pi(x) = \frac{x}{|x|} = \bar{x}$, viz.

$$h_j(\bar{x}) = \nabla \pi(\bar{x}) B_j(\bar{x}) = (\mathrm{Id} - \bar{x} \otimes \bar{x}) B_j \bar{x} = B_j \bar{x} - \langle B_j \bar{x}, \bar{x} \rangle \bar{x} \ , \quad \bar{x} \in \mathbb{S}^{d-1}, \ j = 0, \ldots, m \ . \tag{8.2}$$

Obviously the vector fields are all tangent to the sphere. We combine these vector fields as columns into a single matrix (extended to the whole of $\mathbb{R}^d$) and obtain

$$H(x) = \Big(h_1(x), h_2(x), \cdots, h_m(x)\Big) \in \mathbb{R}^{d\times m} \ , \quad x \in \mathbb{R}^d \ . \tag{8.3}$$

In order to obtain uniqueness of μ we impose the following rank condition which establishes the ellipticity of the angular component. This is a version of the ellipticity condition 12.

Condition 13 (Ellipticity condition)**.** We assume

$$\operatorname{rank}\Big(H^*(\bar{x})\,(A + A_\nu)\,H(\bar{x})\Big) = d - 1 \ , \quad \forall \bar{x} \in \mathbb{S}^{d-1} \ . \tag{8.4}$$

Here A_ν is the positive semidefinite order matrix lower bound in (5.6).

8.1 The projected process. Multiplicative case

We show that the projected process

$$\bar{\xi}_t(x) := \pi(\xi_t(x)) = \frac{\xi_t}{|\xi_t|}(x) \in \mathbb{S}^{d-1} \ , \quad x \in \mathbb{R}^d \setminus \{0\} \ , \tag{8.5}$$

is well defined under the following "no killing" condition that also appears in [AS09] and [LDS02].

Condition 14. The Lévy measure satisfies

$$\nu\Big(z : z^*\boldsymbol{B} = \sum_{j=1}^m z^j B_j = -\operatorname{Id}\Big) = 0 \ . \tag{8.6}$$

This is done by first showing that if the jumps are sufficiently small ξ generates a flow of diffeomorphisms that is sufficiently regular. The result is then extended to the general case via interlacing (*cf.* [AT01, App09]).
Let us first observe the following criterion which shows that Condition 14 is met in our generic situation.

Lemma 8.1. *Assume that the Lévy measure admits a density ϱ with respect to the m'–dimensional Lebesgue measure on some m'–dimensional linear subspace of $\mathbb{R}^m$. Then Condition 14 is met.*

Proof. It is clear that the solution space $\{z : z^*\boldsymbol{B} = -\operatorname{Id}\}$ is a truly affine subspace of $\mathbb{R}^m$ (since it does not contain 0). Recall that ν is assumed to have a density with respect to some linear subspace of $\mathbb{R}^m$. But this implies that the intersection of $\operatorname{supp}\nu$ with any truly affine subspace of $\mathbb{R}^m$ has measure zero. ∎

8.1.1 Small jumps

Recall that by Theorem 2.7 we can restrict the support of the Lévy measure to an environment of 0 such that the matrices $[\mathrm{Id} + z^* \boldsymbol{B}]$ are invertible for ν–a.e. $z \in \mathbb{R}^m$. Indeed in this subsection we assume the following condition.

Condition 15. The support of ν is contained in a ball $\mathbb{B}_\varepsilon$ of radius $\varepsilon > 0$ such that there is a bound $0 < c < 1$ with $\|z^* \boldsymbol{B}\| \leqslant c$ for $z \in \mathbb{B}_\varepsilon$.

Theorem 8.2. *Let ξ be the solution to the linear equation* (7.7) *driven by a Lévy process Z with characteristic triplet (b, A, ν). The projected process $\bar{\xi}_t := \pi(\xi_t)$ is well defined as a process on the unit sphere $\mathbb{S}^{d-1}$. It is given as the unique strong solution to the SDE*

$$\begin{cases} \mathrm{d}\bar{\xi}_t = \bar{X}(\bar{\xi}_t, \mathrm{d}t) \ , \\ \quad \bar{\xi}_0 = \bar{x} \in \mathbb{S}^{d-1} \ . \end{cases} \tag{8.7}$$

Its semimartingale generator is given by

$$\begin{aligned} \bar{X}(x,t) = {} & t\big(h_0(x) + b^* H(x)\big) + (A^{\frac{1}{2}} W_t)^* H(x) \\ & + t \frac{1}{2} \sum_j a_j^2 \big(3\langle B_j x, x\rangle^2 x - 2\langle B_j x, x\rangle B_j x - |B_j x|^2 x\big) \\ & + \iint_0^t \pi\big((\mathrm{Id} + z^* \boldsymbol{B})x\big) - x \ \tilde{N}(\mathrm{d}t\mathrm{d}z) \\ & + t \int \pi\big((\mathrm{Id} + z^* \boldsymbol{B})x\big) - x - z^* H(x) \ \nu(\mathrm{d}z) \ . \end{aligned} \tag{8.8}$$

Proof. By Theorem 2.7 the *SDE* (7.7) generates a flow of diffeomorphisms. This guarantees in particular that $\mathbb{P}\big(\ \xi_t^x = 0 \text{ for some } t \geqslant 0\ \big) = 0$, whenever $x \neq 0$. We can also evoke Theorem 6.4 to conclude that $\mathbb{P}\big(\ |\bar{\xi}_t| = 1,\ \forall t > 0\ \big) = 1$. We apply Itô's formula to the function $\pi(x) = x/|x| = \bar{x}$. (*cf.* [Sko89, eq.(78), p.264]). Observe that for the α-th component of π ($\alpha \in \{1, \dots, d\}$) we have

$$\frac{\partial \pi^\alpha}{\partial x_i}(x) = \frac{\delta_{i\alpha}}{|x|} - \frac{x_i x_\alpha}{|x|^3} \ , \qquad \frac{\partial^2 \pi^\alpha}{\partial x_i \partial x_j}(x) = 3\frac{x_i x_j x_\alpha}{|x|^5} - \frac{1}{|x|^3}\left(x_i \delta_{j\alpha} + x_j \delta_{i\alpha} + x_\alpha \delta_{ij}\right) \ . \tag{8.9}$$

We deduce (8.2) by

$$\nabla\pi(x) B_j(x) = \frac{1}{|x|} B_j x - \frac{1}{|x|^3}\langle B_j x, x\rangle x \tag{8.10}$$

which equals $h_j(x)$ if $|x| = 1$. Let us look at the continuous quadratic covariation part. We have

$$\begin{aligned} & \frac{1}{2}\nabla^2 \pi_{ij}^\alpha(x)\left[B_\ell x \ \mathrm{d}W_t^\ell, B_\ell x \ \mathrm{d}W_t^\ell\right]^{ij} \\ & \qquad = \frac{1}{2}\left\{3\frac{x_i x_j x_\alpha}{|x|^5} - \frac{1}{|x|^3}\left(x_i \delta_{j\alpha} + x_j \delta_{i\alpha} + x_\alpha \delta_{ij}\right)\right\} B_\ell^{ik} B_\ell^{jh} x_k x_h \mathrm{d}t \\ & \qquad = \frac{1}{2}\left\{3\frac{\langle B_\ell x, x\rangle^2}{|x|^5} x_\alpha - \frac{2}{|x|^3}\left(\langle B_\ell x, x\rangle B_\ell x\right)^\alpha - \frac{1}{|x|^3}|B_\ell x|^2 x_\alpha\right\} \mathrm{d}t \ . \end{aligned}$$

Let us turn to the nonlocal part. Clearly the jump kernel is given by $\pi\big((\mathrm{Id}+z^*\boldsymbol{B})x\big) - x$ and the correction term with $|x| = 1$ is given by

$$\begin{aligned} &\mathrm{d}t \int \pi\big((\mathrm{Id}+z^*\boldsymbol{B})x\big) - x - \sum_\ell \nabla\pi(x) z^\ell B_\ell x \; \nu(dz) \\ &= \mathrm{d}t \int \pi\big((\mathrm{Id}+z^*\boldsymbol{B})x\big) - x - \sum_\ell z^\ell h_\ell(x) \; \nu(dz) \; . \end{aligned}$$

Concluding this discussion we obtain the representation

$$\begin{aligned} \mathrm{d}\bar{\xi}_t &= \mathrm{d}t\big(h_0(\bar{\xi}_{t-}) + \sum_j b^j h_j(\bar{\xi}_{t-})\big) + \sum_j a_j h_j(\bar{\xi}_{t-})\mathrm{d}W_t^j \\ &+ t\frac{1}{2}\sum_j a_j^2\big(3\langle B_j\bar{\xi}_{t-},\bar{\xi}_{t-}\rangle^2\bar{\xi}_{t-} - 2\langle B_j\bar{\xi}_{t-},\bar{\xi}_{t-}\rangle B_j\bar{\xi}_{t-} - |B_j\bar{\xi}_{t-}|^2\bar{\xi}_{t-}\big) \\ &+ \iint_0^t \pi\big((\mathrm{Id}+z^*\boldsymbol{B})\bar{\xi}_{t-}\big) - \bar{\xi}_{t-} \;\tilde{N}(\mathrm{d}t\mathrm{d}z) \\ &+ t\int \pi\big((\mathrm{Id}+z^*\boldsymbol{B})\bar{\xi}_{t-}\big) - \bar{\xi}_{t-} - \sum_j z^j h_j(\bar{\xi}_{t-}) \;\; \nu(\mathrm{d}z) \; , \end{aligned} \tag{8.11}$$

which is the explicit form of (8.7). ∎

Theorem 8.3. *The process $\bar{\xi}$ is a strong Feller process on $\mathbb{S}^{d-1}$ and its law possesses a density with respect to the Lebesgue measure on $\mathbb{S}^{d-1}$.*

Proof. We argue with the tools developed in Section 6.2. The jump kernel of (8.8) given by $\gamma(x,z) = \pi\big((\mathrm{Id}+z^*\boldsymbol{B})x\big) - x$ satisfies

$$\gamma'(\bar{x}) = \frac{\partial}{\partial z}\Big|_{z=0}\gamma(\bar{x},z) = H(\bar{x}), \quad \bar{x}\in\mathbb{S}^{d-1} \; . \tag{8.12}$$

Indeed, differentiating the α-th component ($\alpha \in \{1,\dots,d\}$) we obtain

$$\frac{\partial\gamma^\alpha(x,z)}{\partial z^j} = \nabla\pi\big((\mathrm{Id}+z^*\boldsymbol{B})x\big)B_j x \tag{8.13}$$

which by (8.10) reduces to h_j on $\mathbb{S}^{d-1}$ at $z = 0$.
We point out that Condition 13 is the adaption of the ellipticity Condition 12 to the present situation. The theorem then follows by Corollary 6.11 and 6.13. ∎

Corollary 8.4. *$\bar{\xi}$ is uniquely ergodic.*

8.1.2 Interlacing and Large jumps

Having established the regularity for the small–jump process we show that under a mild condition adding large jumps does not destroy the projection with probability 1. To this aim we follow the interlacing strategy (e.g. [App09]). Between the occurrence of large jumps the process is "interlaced" by the process with bounded jumps.
Denote by ξ^ε the solution to (7.7) with the Lévy measure restricted to the ball $\mathbb{B}_\varepsilon$ that satisfies Condition 15. The solution to the full equation is constructed by interlacing.

Proposition 8.5. *(Positivity) Assume that Condition 14 holds. Let $\xi(x)$ be the solution processes to* (7.7) *starting at $x \in \mathbb{R}^d \setminus \{0\}$. Then $\mathbb{P}\big(\ \xi_t(x) = 0\ \big) = 0$ for all $t \geqslant 0$.*

Proof. Let ξ^ε be the solution to (7.7) with Lévy measure satisfying Condition 15, i.e. it is supported inside $\mathbb{B}_\varepsilon$ of the Condition. It generates a flow of diffeomorphisms and hence does not map any $x \neq 0$ to 0 almost surely. Annihilation of the process can therefore only appear at jumps of absolute value greater than ε. These jumps occur as a compound Poisson process with rate $\lambda_\varepsilon = \int_{|z|\geqslant\varepsilon} \nu(\mathrm{d}z)$ and such that ΔZ has the law $\nu_\varepsilon(\mathrm{d}z)$, where $\nu_\varepsilon(\mathrm{d}z) = \lambda_\varepsilon^{-1}\mathbb{1}(|z| \geqslant \varepsilon)\nu(\mathrm{d}z)$. This means that up to an exponential waiting time $\tau \sim \exp(\lambda_\varepsilon)$ the processes ξ and ξ^ε coincide. If at τ the process jumps to zero, then

$$\xi_{\tau-} = \xi^\varepsilon_{\tau-} \in \ker(\mathrm{Id} + \Delta Z^*_\tau \boldsymbol{B}) \tag{8.14}$$

which is a (random) linear subspace of $\mathbb{R}^d$ determined by the first jump of amplitude greater than ε. But since $\bar{\xi}^\varepsilon = \xi^\varepsilon/|\xi^\varepsilon|$ admits a density on $\mathbb{S}^{d-1}$ we deduce that the probability that ξ^ε_τ is in an independent lower dimensional (random) subspace has to be zero. Thus, for the event $\{\xi_\tau = 0\}$ to have positive probability we need $\mathbb{P}\big(\ \Delta Z^*_\tau \boldsymbol{B} = -\mathrm{Id}\ \big)$ to be positive. In fact the two probabilities are equal. We see that the set $\{\xi_t \neq 0, \forall t \geqslant 0\}$ has full probability whenever

$$\nu\Big(z : z^*\boldsymbol{B} = \sum_{j=1}^m z^j B_j = -\mathrm{Id}\Big) = 0\ . \tag{8.15}$$

We can now iterate this argument to later occurrences of large jumps by interlacing. Bear in mind, that at the N-th jump time τ^N, ξ_{τ^N-} admits a density on $\mathbb{S}^{d-1}$ given by a convolution. ∎

Corollary 8.6. *Assume that Condition 14 holds. Then the projected process* (8.5) *is well defined as a process on the unit sphere $\mathbb{S}^{d-1}$. It is given as the unique strong solution to the SDE*

$$\begin{cases} \mathrm{d}\bar{\xi}_t = \bar{X}(\bar{\xi}_t, \mathrm{d}t)\ , \\ \bar{\xi}_0 = \bar{x} \in \mathbb{S}^{d-1}\ . \end{cases} \tag{8.16}$$

Here the semimartingale generator is given by

$$
\begin{aligned}
\bar{X}(x,t) &= t\big(h_0(x) + b^* H(x)\big) + (A^{\frac{1}{2}} W_t)^* H(x) \\
&\quad + t\frac{1}{2}\sum_\ell a_\ell^2 \big(3\langle B_\ell x, x\rangle^2 x - 2\langle B_\ell x, x\rangle B_\ell x - \langle B_\ell x, B_\ell x\rangle x\big) \\
&\quad + \iint_{\mathbb{B}\times[0,t]} \pi\big((\mathrm{Id} + z^* \boldsymbol{B})x\big) - x \ \tilde{N}(\mathrm{d}s\mathrm{d}z) \\
&\quad + \iint_{\mathbb{B}^c\times[0,t]} \pi\big((\mathrm{Id} + z^* \boldsymbol{B})x\big) - x \ N(\mathrm{d}s\mathrm{d}z) \\
&\quad + t\int_{\mathbb{B}} \pi\big((\mathrm{Id} + z^* \boldsymbol{B})x\big) - x - z^* H(x) \ \nu(\mathrm{d}z) \ .
\end{aligned}
\tag{8.17}
$$

8.2 Projection of the canonical equation

Theorem 8.7. *The projected process $\bar{\xi}_t := \pi(\xi_t)$ is well defined as a process on the unit sphere $\mathbb{S}^{d-1}$. It is given as the unique strong solution to the SDE*

$$
\begin{cases}
\mathrm{d}\bar{\xi}_t = \bar{X}^\diamond(\bar{\xi}_t, \mathrm{d}t) \ , \\
\bar{\xi}_0 = \bar{x} \in \mathbb{S}^{d-1} \ .
\end{cases}
\tag{8.18}
$$

The semimartingale generator is given by

$$
\begin{aligned}
\bar{X}^\diamond(x,t) &= t h_0(x) + H(x) \diamond \mathrm{d}Z_t \\
&= t\big(h_0(x) + b^* H(x) + \frac{1}{2}\sum_j a_j^2 h_j^2(x)\big) + (A^{\frac{1}{2}} W_t)^* H(x) \\
&\quad + \iint_{\mathbb{B}\times[0,t]} z^* H(x) \ \tilde{N}(\mathrm{d}t\mathrm{d}z) + \iint_{\mathbb{B}^c\times[0,t]} z^* H(x) \ N(\mathrm{d}t\mathrm{d}z) \\
&\quad + \iint_0^t (\phi^{z^* H}(x) - x - z^* H(x)) \ N(\mathrm{d}t\mathrm{d}z) \ .
\end{aligned}
\tag{8.19}
$$

Proof. We have already argued that the canonical equation generates a stochastic flow of diffeomorphisms such that $\mathbb{P}\big(\ \xi_t \neq 0, \ \forall t > 0 \ \big) = 1$ and the projection to $\mathbb{S}^{d-1}$ is well defined. We can again evoke Theorem 6.4 to conclude that $\mathbb{P}\big(\ |\bar{\xi}_t| = 1, \ \forall t > 0 \ \big) = 1$. We could apply Itô's formula to the projection $\pi(x) = x/|x|$ to obtain (6.14). Instead we can evoke the Leibniz rule (1.35) valid for the canonical equation to deduce

$$
\mathrm{d}[\pi(\xi)]_t = \nabla\pi(\xi) B_0 \xi \mathrm{d}t + \nabla\pi(\xi) B_j \xi \diamond \mathrm{d}Z_t^j = h_0(\bar{\xi})\mathrm{d}t + H(\bar{\xi}) \diamond \mathrm{d}Z_t \ . \tag{8.20}
$$

■

Theorem 8.8. *The process $\bar{\xi}$ is a uniquely ergodic strong Feller process on $\mathbb{S}^{d-1}$ and possesses a transition density p_t.*

Proof. We argue with the tools developed in Section 6.2. By definition the jump kernel of (8.19) given by $\gamma(x,z) = \phi^{z^*H}(x) - x$ is the time 1 flow map generated by the *ODE* $\dot{\phi} = z^*H(\phi)$. Hence

$$\gamma'(x) = \frac{\partial}{\partial z}\Big|_{z=0}\gamma(x,z) = H(x) \ . \tag{8.21}$$

Again Condition 13 is the adaption of Condition 12 to the projected canonical equation. ■

8.3 Furstenberg–Khasminskii averaging

We observe that equation (7.7) generates a flow on $\mathbb{R}^d$ which – as a linear map – can be represented as a matrix valued process Φ solving the matrix valued version of the *SDE*(7.7)

$$\begin{cases} \mathrm{d}\Phi_t = B_0\Phi_t\mathrm{d}t + \displaystyle\sum_{j=1}^{m} B_j\Phi_{t-}(\diamond)\mathrm{d}Z_t^j \ , \\ \Phi_0 = \mathrm{Id} \in \mathbb{R}^{d\times d} \ . \end{cases} \tag{8.22}$$

We actually have $\xi_t^x = \Phi_t x$ for all $t \geqslant 0$, $\mathbb{P}$ almost surely.

Remark 8.9. The flow Φ can be interpreted as a Lévy process in the *monoid* of matrices $\mathbb{R}^{d\times d}$ with the operation of matrix multiplication. If the flow takes values in a matrix or Lie group this point of view can be found e.g. in the monograph [Lia04].

The matrix flow Φ allows us to borrow ergodic results from discrete time *random matrix theory.*

8.3.1 The discrete time (random matrix) setting

We follow the explanations of [Kha11, Chap.6.7], see also [BL85, Chap.1] or [Arn98]. Let $\Phi_1, \Phi_2, \dots \in \mathbb{R}^{d\times d}$ be an *i.i.d.* sequence of invertible $d\times d$ matrices. For some initial $x_0 \in \mathbb{R}^d \setminus 0$ we define iteratively the discrete orbit

$$x_n := \Phi_n x_{n-1} \ , \quad n \in \mathbb{N} \ . \tag{8.23}$$

Then the sequence $x_0, x_1, x_2, \dots$ is also a time-homogeneous Markov chain on $\mathbb{R}^d$. We define

$$\bar{x}_n := \frac{\Phi_n\bar{x}_{n-1}}{|\Phi_n\bar{x}_{n-1}|} \ . \tag{8.24}$$

Then the sequence $\bar{x}_0, \bar{x}_1, \bar{x}_2, \dots$ is also a time-homogeneous Markov chain on the sphere $\mathbb{S}^{d-1}$. For $n \in \mathbb{N}$ let $\lambda_n = \ln|x_n|$ (well defined since $\ker \Phi_n = 0$). Furthermore λ_n satisfies

$$\lambda_n = \lambda_{n-1} + \ln|\Phi_n\bar{x}_{n-1}| = \lambda_0 + \sum_{i=1}^{n} \ln|\Phi_i\bar{x}_{i-1}| \ . \tag{8.25}$$

By compactness of the state space $\mathbb{S}^{d-1}$ it has a stationary probability distribution. We want to apply a Krylov–Bogolyubov type averaging procedure (e.g. [Arn98, Thm. 1.5.8]) and the Birkhoff ergodic theorem to express the limit as an ergodic mean. See also [Fur63, §8].

Theorem 8.10. *Let μ be any invariant probability measure for the Markov chain (x_n) on $\mathbb{S}^{d-1}$ and assume that* $\ln^+ \|\Phi_1\| \in L^1(\Omega, \mathbb{P})$. *Then*

$$\boldsymbol{\lambda} := \lim_{n\to\infty} \frac{\lambda_n}{n} = \mathbb{E}\Big[\int_{\mathbb{S}^{d-1}} \ln|\Phi_1 \bar{x}| \mu(\mathrm{d}\bar{x}) \Big] , \quad \mu\text{-a.s.} \tag{8.26}$$

Proof. Let us first assume that $\ln\|\Phi_1\| \in L^1(\Omega, \mathbb{P})$. Since the matrices $(\Phi_i), i \in \mathbb{N}$, form an *i.i.d.* sequence we may use the canonical model of the probability space as the infinite product

$$\Omega = \Omega_0^{\mathbb{N}} = \big\{ \omega = (\omega_1, \omega_2, \ldots) \ , \ \omega_i \in \Omega_0 \big\} \tag{8.27}$$

of a probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$, endowed with the product σ-algebra and the product measure $\mathbb{P} = \mathbb{P}_0^{\otimes\mathbb{N}}$. The matrices $\Phi_i, i \in \mathbb{N}$, are then given as coordinatewise realizations $\Phi_i(\omega) = \Phi_0(\omega_i)$ of a measurable base functional

$$\Phi_0 : \Omega_0 \to \mathbb{R}^{d\times d} \ . \tag{8.28}$$

The product (*i.i.d.*) structure guarantees that $\mathbb{P}$ is invariant under the canonical shift operator

$$\vartheta\omega = \vartheta(\omega_1, \omega_2, \ldots) = (\omega_2, \omega_3, \ldots) \ . \tag{8.29}$$

The sequences $(x_n)_{n\in\mathbb{N}}$, $(\bar{x}_n)_{n\in\mathbb{N}}$, form *cocycles* over the base shift ϑ in the sense that

$$\bar{x}_{n+m}(\bar{x}_0, \omega) = \bar{x}_n(\bar{x}_m(\bar{x}_0, \omega), \vartheta^m\omega) \ . \tag{8.30}$$

Moreover the $(\lambda_n)_{n\in\mathbb{N}}$ possess the *additive cocycle property* over the *skew product* $\varphi_n := (\bar{x}_n, \vartheta^n)$, $n \in \mathbb{N}$, on $\mathbb{S}^{d-1} \times \Omega$, i.e. for $m, n \in \mathbb{N}, \bar{x}_0 \in \mathbb{S}^{d-1}$ we have

$$\lambda_{m+n}(\bar{x}_0, \omega) = \lambda_m(\bar{x}_0, \omega) + \lambda_n(\bar{x}_m, \vartheta^m\omega) = \lambda_m(\bar{x}_0, \omega) + \lambda_n \circ \varphi_m(\bar{x}_0, \omega) \ . \tag{8.31}$$

Since μ is an invariant measure for the Markov chain $(\bar{x}_n)$ the skew-product structure ensures that $\mu \otimes \mathbb{P}$ is invariant under φ. Thus by virtue of (8.31) and Birkhoff's ergodic theorem (e.g. [Kre85, Thm.2.3., p.9])

$$\boldsymbol{\lambda}(\bar{x}_0, \omega) := \lim_{n\to\infty} \frac{\lambda_n}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \lambda_1 \circ \varphi_k(\bar{x}_0, \omega) \tag{8.32}$$

exists $\mu \otimes \mathbb{P}$-almost surely, is invariant with respect to φ and is constant on the ergodic components of $\mathbb{S}^{d-1} \times \Omega$ with respect to φ. In particular we have

$$\boldsymbol{\lambda}(\bar{x}_0, \omega) \equiv \mathbb{E} \int_{\mathbb{S}^{d-1}} \ln|\Phi_1(\cdot)\bar{x}| \mu(\mathrm{d}\bar{x}) \ , \qquad \text{for } \mu \otimes \mathbb{P}\text{-a.e. } (\bar{x}_0, \omega) \tag{8.33}$$

which is finite by the assumption $\ln\|\Phi_1\| \in L^1(\Omega,\mathbb{P})$. If we only assume that $\ln^+\|\Phi_1\| \in L^1(\Omega,\mathbb{P})$ we localize for $N \in \mathbb{N}$ with

$$\lambda_n^N := \lambda_n \vee (-N) \ . \tag{8.34}$$

The sequence (λ_n^N) is no longer additive but *subadditive*. In this case we should instead replace Birkhoff's theorem by the *subadditive ergodic theorem* of Kingman (Thm. A.1) ([Kre85, Thm.5.3, p.35]) to ensure that the limit in (8.32) exists and satisfies

$$\boldsymbol{\lambda}^N(\bar{x}_0,\omega) \equiv \mathbb{E}\int_{\mathbb{S}^{d-1}} \ln|\Phi_1(\cdot)\bar{x}| \vee (-N)\mu(\mathrm{d}\bar{x}) \ , \qquad \text{for } \mu\otimes\mathbb{P}\text{-a.e. } (\bar{x}_0,\omega) \ . \tag{8.35}$$

We then obtain (8.32) by monotone convergence[1]. ■

8.3.2 The equation in continuous time

Now we return to the question of the existence of the limit

$$\lim_{t\to\infty}\frac{1}{t}\ln|\xi_t(x)| \ , \quad x\in\mathbb{R}^d, \tag{8.36}$$

where ξ is the solution to (7.7). We have seen that the flow generated by (7.7) can be represented by the solution Φ to the matrix valued *SDE* (8.22). We easily deduce from Proposition 7.6 the following result.

Corollary 8.11. *The solution Φ of (7.10) satisfies $\ln^+\|\Phi_t\| \in L^1(\Omega,\mathbb{P})$ for any $t \geqslant 0$.*

We have further seen in Proposition 8.5 that the matrices Φ_t are invertible almost surely. We can then deduce the following from the discrete setup.

Lemma 8.12. *For any time increment $\tau > 0$ we have*

$$\boldsymbol{\lambda}^\tau(x) := \lim_{n\to\infty}\frac{1}{n\tau}\ln|\xi_{n\tau}(x)| \equiv \frac{1}{\tau}\mathbb{E}\Big[\int_{\mathbb{S}^{d-1}}\ln|\Phi_\tau(\cdot)\bar{x}|\mu(\mathrm{d}\bar{x})\ \Big] \ , \ \mu\otimes\mathbb{P}\text{-}a.s.\ . \tag{8.37}$$

Proof. Apply Theorem 8.10 to the sequence of matrices $\Phi_{n\tau}\Phi^{-1}_{(n-1)\tau}$ which are *i.i.d.* copies of Φ_τ. ■

We are now in the position to obtain Furstenberg–Khasminskii's formula in continuous time by sending τ to zero. Formally the argument proceeds as follows

$$\frac{1}{\tau}\mathbb{E}\big[\ \ln|\Phi_\tau(\cdot)\bar{x}|\ \big] = \frac{1}{2\tau}\mathbb{E}\big[\ \ln|\xi_\tau(\bar{x})|^2\ \big] = \frac{1}{2}\big(\frac{1}{\tau}\mathcal{P}_\tau\ln|\cdot|^2\big)(\bar{x}) \xrightarrow{\tau\searrow 0} \frac{1}{2}\mathcal{L}^{(\diamond)}\big(\ln|\cdot|^2\big)(\bar{x}) \ .$$

However the limit on the right hand side has to be handled with care.

[1] To be technically precise we apply Lebesgue's monotone convergence theorem on the $\mu\otimes\mathbb{P}$–measurable set where $\ln|\Phi_1(\omega)\bar{x}| \leqslant 0$.

Theorem 8.13. *Let ξ be the Markov process solving the stochastic differential equation (7.7) and $\mathcal{L}^{(\diamond)}$ the generator of the associated semigroup. Then the top Lyapunov exponent is given by the formula*

$$\boldsymbol{\lambda} := \lim_{t\to\infty} \frac{1}{t} \ln |\xi_t(x)| = \int_{\mathbb{S}^{d-1}} \frac{1}{2} \mathcal{L}^{(\diamond)}\big(\ln|\cdot|^2\big)(\bar{x})\mu(\mathrm{d}\bar{x}) \ , \tag{8.38}$$

where μ is the unique ergodic measure on $\mathbb{S}^{d-1}$ of the projected process $\bar{\xi} = \xi/|\xi|$.
The integrand is the infinitesimal generator $\mathcal{L}^{(\diamond)}$ of ξ acting on the logarithm of norm squared. It takes the form

(i) in the case of the multiplicative jump kernel

$$\begin{aligned} \mathcal{L}\big(\ln|\cdot|^2\big)(\bar{x}) = {} & 2\left(q_0(\bar{x}) + b^j q_j(\bar{x})\right) \\ & + a^j\left(|B_j\bar{x}|^2 - 2\big(\bar{x}^* B_j \bar{x}\big)^2\right) \\ & + 2\int_{\mathbb{R}^m} \log(|\bar{x} + z^*\boldsymbol{B}\bar{x}|) - \mathbb{1}_{\{|z|\leqslant 1\}} z^j q_j(\bar{x}) \ \nu(\mathrm{d}z) \ , \end{aligned} \tag{8.39}$$

(ii) in the case of the canonical jump kernel

$$\begin{aligned} \mathcal{L}^{\diamond}\big(\ln|\cdot|^2\big)(\bar{x}) = {} & 2(q_0(\bar{x}) + b^j q_j(\bar{x})) + a^j h_j(q_j)(\bar{x}) \\ & + 2\int_{\mathbb{R}^m}\int_0^1 z^j q_j(\phi_r^{z^*\boldsymbol{B}}\bar{x})\mathrm{d}r - \mathbb{1}_{\{|z|\leqslant 1\}}\left(\bar{x}^*\phi^{z^*\boldsymbol{B}}\bar{x} - 1\right)\mathrm{d}r\nu(\mathrm{d}z) \end{aligned} \tag{8.40}$$

with the following notation (e.g. cf. [Arn98, p.254]) for $\bar{x} \in \mathbb{S}^{d-1}$

$$\begin{aligned} q_j(\bar{x}) &= \langle B_j\bar{x}, \bar{x}\rangle, \\ h_j(\bar{x}) &= B_j\bar{x} - q_j(\bar{x})\bar{x}, \\ h^j(q_j)(\bar{x}) &= h_j^i \frac{\partial q_j}{\partial x_i}(\bar{x}) = \nabla q_j(\bar{x}) h_j(\bar{x}), \\ &= \langle (B_j + B_j^*)\bar{x}, B_j\bar{x}\rangle - 2\langle B_j\bar{x}, \bar{x}\rangle^2 \ . \end{aligned} \tag{8.41}$$

Remark 8.14. If there are no jumps and no drift ($\nu = 0, b = 0$), then (8.40) coincides with the representation for linear Stratonovich equations and is well known *cf.* equation (6.2.20) in [Arn98, p.255].

Proof. Similarly to the proof of Proposition 7.6 we define a sequence of smooth localization functions $\chi^N : \mathbb{R}^+ \to [0,1], N \in \mathbb{N}$ satisfying

$$\begin{aligned} \chi^N(x) &= 0, \quad 0 < x \leqslant \mathrm{e}^{-N} \\ \chi^N(x) &= 1, \quad x \geqslant \mathrm{e}^{-(N-1)} \end{aligned}$$

Then we have

$$\chi^N \times \ln(x) \leqslant \chi^{N-1} \times \ln(x) \leqslant \cdots \leqslant \chi^1 \times \ln(x) \leqslant \ln^+(x)$$

and $\ln^+(x) \leqslant |\chi^N(x)\ln(x)| \leqslant N \vee \ln^+(x)$. By the same arguments as in the proof of Proposition 7.6 (the function χ there corresponds to χ^1) we see that

$$\frac{1}{2}(\chi^N \times \ln|\cdot|^2)(\xi_t(x)) - \int_0^t \frac{1}{2}\mathcal{L}^{(\diamond)}\left(\chi^N \times \ln|\cdot|^2\right)(\xi_s(x))\mathrm{d}s \tag{8.42}$$

is a true martingale by Dynkin's formula. And hence from (8.37)

$$\boldsymbol{\lambda}^\tau(x) \leqslant \mathbb{E}\Big[\int_{\mathbb{S}^{d-1}} \frac{1}{\tau}\int_0^\tau \frac{1}{2}\mathcal{L}^{(\diamond)}\left(\chi^N \times \ln|\cdot|^2\right)(\xi_s(\bar{x}))\mathrm{d}s\ \mu(\mathrm{d}\bar{x}) \Big] \tag{8.43}$$

Sending first $\tau \searrow 0$ and then $N \nearrow \infty$ we obtain with the monotone convergence theorem of Lebesgue

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}^0(x) = \mathbb{E}\Big[\int_{\mathbb{S}^{d-1}} \frac{1}{2}\mathcal{L}^{(\diamond)}\left(\ln|\cdot|^2\right)(\bar{x})\ \mu(\mathrm{d}\bar{x}) \Big]\ , \tag{8.44}$$

where the integrand neither depends on ω nor on x.
It remains to express the action of the generator on the logarithm. We investigate the multiplicative case first.

(i) The generator $\mathcal{L}$ is given in (7.12). Let us first notice that

$$(\nabla \ln|\cdot|^2)(x)B_jx = q_j(\bar{x})\ ,$$
$$(\langle B_jx, \nabla\rangle \ln|\cdot|^2)^2(x) = |B_j\bar{x}|^2 - 2\big(\bar{x}^*B_j\bar{x}\big)^2\ .$$

Similarly the compensator in the non-local part is

$$(\nabla \ln|\cdot|^2)(x)z^jB_jx = z^jq_j(\bar{x})\ .$$

(ii) In the canonical interpretation $\mathcal{L}^\diamond$ is given in (7.25). We concentrate on the non-local part. It satisfies

$$\begin{aligned}
&\int \ln|\phi^{z^*\boldsymbol{B}}x|^2 - \ln|x|^2 - \mathbb{1}_{\{|z|\leqslant 1\}}\nabla\ln|\cdot|^2\left(\phi^{z^*\boldsymbol{B}}x - x\right)\nu(\mathrm{d}z)\\
&\quad= \int\int_0^1 \nabla\ln|\cdot|^2(\phi_r^{z^*\boldsymbol{B}}x)z^*\boldsymbol{B}\phi_r^{z^*\boldsymbol{B}}x\ \mathrm{d}r - \mathbb{1}_{\{|z|\leqslant 1\}}\frac{x^*}{|x|^2}\left(\phi^{z^*\boldsymbol{B}}x - x\right)\mathrm{d}r\nu(\mathrm{d}z)\\
&\quad= \int\int_0^1 z^jq_j(\phi_r^{z^*\boldsymbol{B}}\bar{x})\ \mathrm{d}r - \mathbb{1}_{\{|z|\leqslant 1\}}\left(\bar{x}^*\phi^{z^*\boldsymbol{B}}\bar{x} - 1\right)\mathrm{d}r\nu(\mathrm{d}z)\ .
\end{aligned}$$

■

Appendices

Appendix A

Some ergodic theory

A.1 Kingman's subadditive ergodic theorem

For convenience we state a version of Kingman's subadditive ergodic theorem here following [GM89, Thm.1.1] (see also [Arn98] or [Kre85]).

Theorem A.1 (Kingman)**.** *Let $f_n : \Omega \to \mathbb{R} \cup \{-\infty\}$ be a sequence of measurable functions with $f_1^+ \in L^1(\Omega)$. Let $(f_n)_{n\in\mathbb{N}}$ be* SUBADDITIVE *viz.*

$$f_{m+n} \leqslant f_m(\omega) + f_n(\vartheta^m \omega) \quad \mathbb{P}\text{-}a.s. \tag{A.1}$$

Then there is a ϑ-invariant measurable function $f : \Omega \to \mathbb{R} \cup \{-\infty\}$ with $f^+ \in L^1(\Omega)$ such that $\lim_{n\to\infty} n^{-1} f_n(\omega) = f(\omega)$ for almost all $\omega \in \Omega$, and

$$\lim_{n\to\infty} \frac{1}{n} \int_\Omega f_n(\omega)\mathbb{P}(\mathrm{d}\omega) = \int_\Omega f(\omega)\mathbb{P}(\mathrm{d}\omega) \ . \tag{A.2}$$

Appendix B

More on stochastic integration

B.1 L^p-estimates and Itô's formula

For the convenience of the reader we give here some standard results of stochastic analysis for jump diffusions used in this work. Throughout this section ξ is the unique strong solution to an *SDE*

$$\begin{cases} \mathrm{d}\xi_t = X(\xi_{t-}, \mathrm{d}t) \ , \\ \xi_0 = x \in \mathbb{R}^d \ , \end{cases} \tag{B.1}$$

for a semimartingale generator X to be specified but such that a unique strong solution exists for all $t \geqslant 0$ (e.g. Condition 1).

B.1.1 Itô's formula for *SDE*

We state a version of Itô's formula for solutions of *SDE* (See e.g. [IW89, Chap.II, Thm.5.1., p.66f]). Let the semimartingale generator of (B.1) be of the form

$$\begin{aligned} X(x,t) = \beta(x)t + \sigma(x)W_t + \iint_{\mathbb{B}\times[0,t]} \gamma(x,z)\tilde{N}(\mathrm{d}z\mathrm{d}s) \\ + \iint_{\mathbb{B}^c\times[0,t]} \gamma(x,z)N(\mathrm{d}z\mathrm{d}s) \ , \end{aligned} \tag{B.2}$$

with parameters β, σ and γ as in 1.2. Itô's formula then reads as follows.

Proposition B.1 (Itô's formula). *Let $F \in \mathscr{C}^2(\mathbb{R}^d)$. Then the process $F(\xi_t)$ satisfies*

$$\begin{aligned} \mathrm{d}F(\xi_t) = & \nabla F(\xi_t)\beta(\xi_t)\mathrm{d}t + \nabla F(\xi_t)\sigma(\xi_t)\mathrm{d}W_t \\ & + \frac{1}{2}\operatorname{Tr}[\sigma^*\nabla^2 F\sigma](\xi_t)\mathrm{d}t \\ & + \int_{\mathbb{B}} F(\xi_{t-} + \gamma(\xi_{t-}, z)) - F(\xi_{t-} + \gamma(\xi_{t-}, z))\tilde{N}(\mathrm{d}z\mathrm{d}t) \\ & + \int_{\mathbb{B}^c} F(\xi_{t-} + \gamma(\xi_{t-}, z)) - F(\xi_{t-} + \gamma(\xi_{t-}, z))N(\mathrm{d}z\mathrm{d}t) \\ & + \mathrm{d}t\int_{\mathbb{B}^c} F(\xi_{t-} + \gamma(\xi_{t-}, z)) - F(\xi_{t-} + \gamma(\xi_{t-}, z)) - \nabla F(\xi_{t-})\gamma(\xi_{t-}, z)\nu(\mathrm{d}z) \end{aligned} \tag{B.3}$$

B.1.2 L^p-estimates for *SDE*

Let us note the following L^p-estimates (*cf.* [Kun04, Thm.2.11]).

Proposition B.2 (L^p-estimates). *Let X be given by the semimartingale field*

$$X(x,t) = \beta(x)t + \sigma(x)W_t + \iint_{\mathbb{B}\times[0,t]} \gamma(x,z)\tilde{N}(\mathrm{d}z\mathrm{d}s)\ . \tag{B.4}$$

and ξ the solution to (1.14)*. Then the we have*

$$\begin{aligned} \mathbb{E}\Big[\ \sup_{s<t}|\xi_t|^p\ \Big] \lesssim_p & \ \mathbb{E}\Big[\ \Big(\int_0^t |\beta(\xi_s)|\mathrm{d}s\Big)^p\ \Big] + \mathbb{E}\Big[\ \Big(\int_0^t |\sigma(\xi_s)|^2\mathrm{d}s\Big)^{p/2}\ \Big] \\ & + \mathbb{E}\Big[\ \Big(\iint_{\mathbb{B}\times[0,t]} |\gamma(\xi_s,z)|^2\mathrm{d}s\nu(\mathrm{d}z)\Big)^{p/2}\ \Big] \\ & + \mathbb{E}\Big[\ \iint_{\mathbb{B}\times[0,t]} |\gamma(\xi_s,z)|^p\mathrm{d}s\nu(\mathrm{d}z)\ \Big]\ . \end{aligned}$$

B.2 Graded *SDE*

Not all *SDE* considered in this thesis fulfill the Lipschitz condition 1 (e.g. the Jacobian in Section 2.2.1). Thus existence and uniqueness of a solution cannot be guaranteed by Theorem 1.3. Instead we rely on the notion of *graded systems* (see e.g. [Str81a] in a similar context).

A graded system is a system of coupled *SDE* where the coefficients have an "lower triangular" Lipschitz structure. Such a system allows for the decomposition of the state space $\mathbb{R}^d$ into subspaces representing a hierarchy of dependence such that on each subspace the equation may not be Lipschitz in the depended variables but need to be Lipschitz with respect to the independent variables. Existence and uniqueness is then established by an induction argument. This philosophy is made precise in the following definition.

Definition B.3 (graded system)**.** A GRADED SYSTEM is an *SDE* (1.14) where the semimartingale generator X in (1.15) has the following lower triangular structure. There is a partition $d^1 + \cdots + d^k = d$ of d such that for $x = (x^1, \ldots, x^k) \in \mathbb{R}^{d^1} \times \cdots \times \mathbb{R}^{d^2} = \mathbb{R}^d$ we have

$$X(x,t) = \begin{pmatrix} X^1(x^1,t) \\ X^2(x^1,x^2,t) \\ \vdots \\ X^k(x^1,\ldots,x^k,t) \end{pmatrix} \tag{B.5}$$

where every X^j is a $\mathbb{R}^{d^j}$-valued semimartingale generator of the form (1.16).

Theorem B.4 (existence and uniqueness II)**.** *Assume that the semimartingale generator X admits a grading according to* (B.5) *such that for any $j \in 1, \ldots, k$*

$$X^j(x^1, x^2, \ldots, x^j) \textit{ is Lipschitz in } x^j \ ,$$

and Hölder in $x^1, \ldots, x^{j-1}$. Then the equation (1.21) *has a unique solution in L^p.*

Bibliography

[AD09] Frank Aurzada and Steffen Dereich. Small deviations of general Lévy processes. *Annals of Probability*, 37(5):2066–2092, 2009.

[AK93] David Applebaum and Hiroshi Kunita. Lévy flows on manifolds and Lévy processes on Lie groups. *Journal of Mathematics of Kyoto University*, 33(4):1103–1123, 1993.

[App09] David Applebaum. *Lévy Processes and Stochastic Calculus*, volume 116 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, 2nd edition, 2009.

[Arn98] Ludwig Arnold. *Random Dynamical Systems.* Springer Monographs in Mathematics. Springer-Verlag, 1998.

[AS09] David Applebaum and Michailina Siakalli. Asymptotic Stability of Stochastic Differential Equations driven by Lévy noise. *Journal of Applied Probabiliy*, 46(4):1116–1129, 2009.

[AS10] David Applebaum and Michailina Siakalli. Stochastic stabilization of dynamical systems using Lévy noise. *Stochastics and Dynamics*, 10(04):509–527, 2010.

[AT01] David Applebaum and Fuchang Tang. Stochastic flows of diffeomorphisms on manifolds driven by infinite-dimensional semimartingales with jumps. *Stochastic Processes and their Applications*, 92:219–236, 2001.

[BC86] Richard F. Bass and Michael Cranston. The Malliavin calculus for pure jump processes and applications to local time. *The Annals of Probability*, 14(2):490–532, 1986.

[BDV05] Christian Bonatti, Lorenzo J. Díaz, and Marcelo Viana. *Dynamics Beyond Uniform Hyperbolicity. A Global Geometric and Probabilistic Perspective*, volume 102 of *Encyclopedia of Mathematical Sciences.* Springer-Verlag, 2005.

[Bel87] Denis Bell. *The Malliavin Calculus*, volume 34 of *Pitman Monographs and Surveys in Pure and Applied Mathematics.* Longman Scientific & Technical, 1987.

[Ber96] Jean Bertoin. *Lévy processes.* Cambridge University Press, 1996.

[BG80] Richard L. Bishop and Samuel I. Goldberg. *Tensor Analysis on Manifolds.* Dover Publications, Inc., 1980.

[BGJ87] Klaus Bichteler, Jean-Bernard Gravereaux, and Jean Jacod. *Malliavin calculus for processes with jumps.* Stochastics Monographs. Gordon and Breach Science Publishers, 1987.

[BGT87] Nicholas H. Bingham, Charles M. Goldie, and Jef L. Teugels. *Regular variation*, volume 2 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, 2^{nd} edition, 1987.

[Bil99] Patrick Billingsley. *Convergence of Probability Measures.* Wiley series in probability and statistics. John Wiley & Sons, 2^{nd} edition, 1999.

[Bis81] Jean-Michel Bismut. Martingales, the Malliavin Calculus and Hypoellipticity Under General Hörmander's Conditions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 56:469–505, 1981.

[BL85] Philippe Bougerol and Jean Lacroix. *Products of Random Matrices with Applications to Schrödinger Operators*, volume 8 of *Progress in Probability and Statistics.* Birkäuser, 1985.

[BL86a] G. L. Blankenship and C. W. Li. Almost sure stability of linear stochastic systems with Poisson process coefficients. *SIAM Journal on Applied Mathematics*, 46(5):875–911, 1986.

[BL86b] G. L. Blankenship and Kenneth A. Loparo. Almost Sure Instability of a Class of Linear Stochastic Systems with Jump Coefficitents. In Ludwig Arnold and Volker Wihstutz, editors, *Lyapunov Exponents, Lecture Notes in Mathematics Volume 1186*, pages 160–190. Springer, 1986.

[Bog10] Vladimir I. Bogachev. *Differentiable Measures and the Malliavin Calculus*, volume 164 of *Mathematical Surveys and Monographs.* American Mathematical Society, 2010.

[CDFG81] Noel Cressie, Anne Davis, Leroy Folks, and Policello. George. The Moment-Generating Function and Negative Integer Moments. *The American Statistician*, 35(3):148–150, 1981.

[CF07] Thomas R. Cass and Peter K. Friz. The Bismut-Elworthy-Li formula for jump-diffusions and applications to Monte Carlo pricing in finance (Working Paper). *ArXiv e-prints*, pages 1–30, 2007, 0604311.

[CFT13] Oswaldo Luiz do Valle Costa, Marcelo D. Fragoso, and Marcos G. Todorov. *Continuous-Time Markov Jump Linear Systems.* Probability and Its Applications. Springer-Verlag, 2013.

[CH15] Bertrand Cloez and Martin Hairer. Exponential ergodicity for Markov processes with random switching. *Bernoulli*, 21(1):505–536, 2015.

[Çin11] Erhan Çinlar. *Probability and Stochastics*, volume 261 of *Graduate Texts in Mathematics.* Springer-Verlag, 2011.

[EL94] K. David Elworthy and Xue-Mei Li. Formulae for the Derivatives of Heat Semigroup. *Journal of Functional Analysis*, 125:252–286, 1994.

[FK85] Tsukasa Fujiwara and Hiroshi Kunita. Stochastic differential equations of jump type and Lévy processes in diffeomorphisms group. *Kyoto Journal of Mathematics*, 1(25):71–106, 1985.

[FK99] Tsukasa Fujiwara and Hiroshi Kunita. Canonical SDE's based on semmimartingales with spatial parameters; Part I: Stochastic flows of diffeomorphisms. *Kyushu Journal of Mathematics*, 53:265–300, 1999.

[FTT14] Damir Filipović, Stefan Tappe, and Josef Teichmann. Invariant manifolds with boundary for jump-diffusions. *Electronic Journal of Probability*, 19, jun 2014.

[Fuj91] Tsukasa Fujiwara. Stochastic differential equations of jump type on manifolds and Lévy flows. *Kyoto Journal of Mathematics*, 1(31):99–119, 1991.

[Fur63] Hillel Furstenberg. Noncommuting random products. *Transactions of the American Mathematical Society*, 108(377-428), 1963.

[GM89] I. Ya. Gol'dsheid and G. A. Margulis. Lyapunov indices of a product of random matrices. *Russian Mathematical Surveys*, 44(1):11–71, 1989.

[Hai08] Martin Hairer. Ergodic theory for Stochastic PDEs. *Lecture Notes at the University of Warwick (available at http://hairer.org)*, 2008.

[IK06] Yasushi Ishikawa and Hiroshi Kunita. Malliavin calculus on the Wiener–Poisson space and its application to canonical SDE with jumps. *Stochastic Processes and their Applications*, 116(12):1743–1769, 2006.

[IKM15] Dmytro Ivanenko, Alexei M. Kulik, and Hiroki Masuda. Uniform LAN property of locally stable Lévy process observed at high frequency. *ALEA, Lat. Am. J. Probab. Stat.*, 12(2):835–862, 2015.

[Ish13] Yasushi Ishikawa. *Stochastic Calculus of Variations for Jump Processes*, volume 54 of *De Gruyter Studies in Mathematics.* Walter de Gruyter, 2013.

[IW89] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland/Kodansha, 2nd edition, 1989.

[Jac82] Jean Jacod. Équations differentielles stochastiques lineaires: La méthode de variation des constantes. In *Séminaire de Probabilités XVI 1980/81*, volume 920 of *Lecture Notes in Mathematics*, pages 442–446. Springer-Verlag, 1982.

[JS03] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes.* Grundlehren der mathematischen Wissenschaften 288. Springer-Verlag, softcover reprint of the hardcover 2nd edition, 2003.

[Kha67] Rafail Khasminskii. Necessary and sufficient conditions for the asymptotic stability of linear stochastic systems. *Theory of Probability & Its Applications*, 12(1):144–147, 1967.

[Kha11] Rafail Khasminskii. *Stochastic Stability of Differential Equations.* Applications of Mathematics. Springer-Verlag, 2nd edition, 2011.

[KPP95] Thomas G. Kurtz, Étienne Pardoux, and Philipp Protter. Stratonovich stochastic differential equations driven by general semimartingales. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, 32(2):351–377, 1995.

[Kre85] Ulrich Krengel. *Ergodic Theorems*, volume 6 of *De Gruyter Studies in Mathematics.* Walter de Gruyter, 1985.

[KT01] Takashi Komatsu and Atsushi Takeuchi. On the Smoothness of PDF of Solutions to SDE of Jump Type. *International Journal of Differential Equations and Applications*, 2(2):141–197, 2001.

[Kun90] Hiroshi Kunita. *Stochastic Flows and Stochastic Differential Equations.* Cambridge University Press, 1990.

[Kun95] Hiroshi Kunita. Some problems concerning Lévy processes on Lie groups. In Michael C. Cranston and Mark A. Pinsky, editors, *Stochastic analysis. Proceedings of the Summer Research Institute on stochastic analysis, held at Cornell University, Ithaca, NY, USA, July 11-30, 1993.*, volume 57 of *Proceedings of Symposia in Pure Mathematics*, pages 323–341. American Mathematical Society, 1995.

[Kun99] Hiroshi Kunita. Canonical Stochastic Differential Equations based on Lévy processes and their Supports. In Hans Crauel and Matthias Gundlach, editors, *Stochastic Dynamics*, chapter 12, pages 283–304. Springer-Verlag, 1999.

[Kun01] Hiroshi Kunita. Malliavin calculus of canonical stochastic differential equations with jumps. アカデミア 南山大学紀要 数理情報編, 1:305–373, 2001.

[Kun04] Hiroshi Kunita. Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In M.M. Rao, editor, *Real and Stochastic Analysis: New Perspectives*, pages 305–373. Boston, MA: Birkhäuser, 2004.

[Kun11] Hiroshi Kunita. Analysis of nondegenerate Wiener-Poisson functionals and its applications to Itô's SDE with jumps. *Sankhyā: The Indian Journal of Statistics*, 73(A):1–45, 2011.

[KZY07] R. Z. Khasminskii, C. Zhu, and G. Yin. Stability of regime-switching diffusions. *Stochastic Processes and their Applications*, 117(8):1037–1051, 2007.

[LDS02] C.W. Li, Z. Dong, and R. Situ. Almost sure stability of linear stochastic differential equations with jumps. *Probability Theory and Related Fields*, 123:121–155, 2002.

[Lia04] Ming Liao. *Lévy Processes in Lie Groups*. Cambridge Tracts in Mathematics. Cambridge University Press, 2004.

[LMR14] Sean D. Lawley, Jonathan C. Mattingly, and Michael C. Reed. Sensitivity to switching rates in stochastically switched odes. *Communications in Mathematical Sciences*, 12(7):1343–1352, 2014.

[Lya92] Aleksandr Mikhailovich Lyapunov. Stability of motion: general problem. *International Journal of Control*, 55(3):767–772, 1992.

[Mal97] Paul Malliavin. *Stochastic Analysis*, volume 313 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 1997.

[Mao94] Xuerong Mao. *Exponential Stability of Stochastic Differential Equations*. Pure and Applied Mathematics. Marcel Dekker Inc., 1994.

[Mao99] Xuerong Mao. Stability of stochastic differential equations with Markovian switching. *Stochastic Processes and their Applications*, 79:45–67, 1999.

[Mar81] Steven I. Marcus. Modeling and approximation of stochastic differential equations driven by semimartingales. *Stochastics and Stochastics Reports*, 4:233–245, 1981.

[MRA07] Jerrold E. Marsden, Tudor Ratiu, and Ralph Abraham. *Manifolds, Tensor Analysis, and Applications.* preprint. Springer-Verlag Berlin Heidelberg, 3$^{\text{rd}}$ edition, 2007.

[MS96] Salah-Eldin Mohammed and Michael Scheutzow. Lyapunov exponents of linear stochastic functional equations driven by semimartingales. Part I: the multiplicative ergodic theory. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, 32(1):69–105, 1996.

[MS97] Salah-Eldin Mohammed and Michael Scheutzow. Lyapunov exponents of linear stochastic functional-differential equations. II. Examples and case studies. *The Annals of Probability*, 25(3):1210–1240, 1997.

[Nor88] James R. Norris. Integration By Parts For Jump Processes. In *Séminaire de Probabilités XXII*, volume 1321 of *Lecture Notes in Mathematics*, pages 271–315. Springer-Verlag, 1988.

[Nua06] David Nualart. *The Malliavin Calculus and Related Topics*. Probability and Its Applications. Springer-Verlag, 2nd edition, 2006.

[Oh00] Jae-Pill Oh. The density for jump processes in canonical stochastic differential equation. *Kangweon-Kyungki Mathematical Journal*, 8(1):1–17, 2000.

[Ose68] Valery I. Oseledec. A multiplicative ergodic theorem. Ljapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, 19:197–231, 1968.

[Pit63] T. S. Pitcher. Likelihood ratios for stochastic processes related by groups of transformations. *Illinois J. Math.*, 7(3):396–414, 1963.

[Pit64] T. S. Pitcher. Likelihood ratios for stochastic processes related by groups of transformations II. *Illinois J. Math.*, 8(2):271–279., 1964.

[Pro04] Philip E. Protter. *Stochastic integration and differential equations*. Springer-Verlag, Berlin Heidelberg, 2004.

[Sat99] Ken-iti Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1999.

[Sat11] Ken-iti Sato. レヴィ過程による確率積分と無限分解可能分布. *Sūgaku*, 63(2):161–181, 2011.

[Sat14] Ken-iti Sato. Stochastic integrals with respect to Lévy processes and infinitely divisible distributions. *Sūgaku Expositions*, 27(1):161–181, 2014.

[Sim04] Thomas Simon. Small ball estimates in p-variation for stable processes. *Journal of Theoretical Probability*, 17(4):979–1002, 2004, 0306015v1.

[Sko89] Anatoliy V. Skorokhod. *Asymptotic Methods in the Theory of Stochastic Differential Equations*, volume 78 of *Translation of Mathematical Monographs*. American Mathematical Society, 1989.

[Son14] Yu Lin Song. Gradient estimates and coupling property for semilinear SDEs driven by jump processes. *Science in China Series A: Mathematics*, 2014.

[Str81a] Daniel W. Stroock. The Malliavin Calculus, a Functional Analytic Approach. *Journal of Functional Analysis*, 44(198 1):212–257, 1981.

[Str81b] Daniel W. Stroock. The Malliavin calculus and its application to second order parabolic differential equations: Part I. *Mathematical Systems Theory*, 14:25–65, 1981.

[Tak10] Atsushi Takeuchi. Bismut–Elworthy–Li-Type Formulae for Stochastic Differential Equations with Jumps. *Journal of Theoretical Probability*, 23(2):576–604, February 2010.

[Via14] Marcelo Viana. *Lectures on Lyapunov Exponents*, volume 145 of *Cambridge studies in advances mathematics*. Cambridge University Press, 2014.

[Wat84] Shinzo Watanabe. *Lectures on Stochastic Differential Equations and Malliavin Calculus*. Tata Institute of Fundamental Research. Springer-Verlag, 1984.

[WXZ15] Feng Yu Wang, Lihu Xu, and Xicheng Zhang. Gradient estimates for SDEs driven by multiplicative Lévy noise. *Journal of Functional Analysis*, 269(10):3195–3219, 2015, 1301.4528.

[WZ65] Eugene Wong and Moshe Zakai. On the convergence of ordinary integrals to stochastic integrals. *The Annals of Mathematical Statistics*, 36(5):1560–1564, 1965.

[Yos80] Kôsaku Yosida. *Functional Analysis*, volume 123 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 6^{th} edition, 1980.

[YY12] Zhixin Yang and G. Yin. Stability of nonlinear regime-switching jump diffusion. *Nonlinear Analysis, Theory, Methods and Applications*, 75(9):3854–3873, 2012.

[Zei86] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications. I: Fixed Point Theorems*. Springer-Verlag, 1986.

Notation

Spaces (of functions and measures)

$\mathbb{R}_0^d$	the non-zero real vectors ($\mathbb{R}_0^d = \mathbb{R}^d \setminus \{0\}$)
$\mathbb{R}^{d*}$	the dual of $\mathbb{R}^d$ identified as d-dimensional row vectors ($\mathbb{R}^{d*} \simeq \mathbb{R}^d$)
$\mathbb{S}^{d-1}$	the $d-1$-dimensional unit sphere (as submanifold of $\mathbb{R}^d$)
$T_{\bar{x}}\mathbb{S}^{d-1}$	the $d-1$-dimensional tangent plane to the unit sphere at $\bar{x} \in \mathbb{S}^{d-1}$
$\mathbb{B}, \mathbb{B}_\varepsilon$	the unit ball, resp. ball of radius $\varepsilon > 0$ (usually in $\mathbb{R}^m$)
$\mathbb{M}$	a smooth submanifold of $\mathbb{R}^d$
$\mathbb{X}$	a Banach space, usually complete and separable (i.e. "Polish")
$\mathcal{B}(\mathbb{X})$	the set of all Borel sets on $\mathbb{X}$
$\mathscr{M}(\mathbb{X})$	the Radon measures on $\mathbb{X}$
$\mathscr{M}_{\mathcal{B}}(\mathbb{X})$	the Borel measures on $\mathbb{X}$
$\mathscr{P}(\mathbb{X})$	the probability measures on $\mathbb{X}$
$\mathscr{P}_{\mathbb{P}}(\Omega \times \mathbb{X})$	the probability measures on $\Omega \times \mathbb{X}$ with fixed marginal $\mathbb{P}$ on Ω
$\mathscr{B}(\mathbb{X})$	the Borel measurable functions on $\mathbb{X}$
$\mathscr{B}_b(\mathbb{X})$	the bounded Borel measurable functions on $\mathbb{X}$
$\mathscr{C}, \mathscr{C}^k(\mathbb{X})$	the real valued continuous (n-times continuously differentiable) functions on $\mathbb{X}$
$\mathscr{C}_b, \mathscr{C}_c, \mathscr{C}_\infty(\mathbb{X})$	the subspace of functions with are bounded / compactly supported /
$\mathscr{C}_b^k$	functions with all (partial) derivatives up to order k bounded
$L^p(\mathbb{X}, \mu)$	the space (of equivalence classes) of p-integrable functions
$\mathbb{D}([0,T];\mathbb{X})$	the Skorokhod space of càdlàg functions from the interval $[0,T]$ to the metric space $\mathbb{X}$
$\mathbb{D}_0([0,T];\mathbb{X})$	the subset of Skorokhod space $\mathbb{D}([0,T];\mathbb{X})$ of functions starting at 0
π	denotes the projection to the unit sphere $\pi : x \mapsto \bar{x} = \frac{x}{\lvert x \rvert}$

Norms

$\lvert \cdot \rvert$	Euclidean norm of vectors and scalars (tensors of rank $\leqslant 1$)
$\lVert \cdot \rVert$	the (induced operator) norm of tensors of rank $\geqslant 2$, (matrices, functions)
$\lvert \cdot \rvert_{[0,T]}$	is the supremum norm with respect to time $\lvert X \rvert_{[0,T]} = \sup_{t \in [0,T]} \lvert X(t, \cdot) \rvert$
$\lVert \cdot \rVert_{[0,T]}$	denotes the supremum operator norm with respect to time $\lVert F \rVert_{[0,T]} = \sup_{t \in [0,T]} \lVert F(t, \cdot) \rVert$.

Stochastic analysis

$\diamond \mathrm{d}Z, \circ \mathrm{d}W$	Marcus canonical integration wrt. Z or Stratonovich integration wrt. W

Binary relations

$x \perp y$	the vectors are x, y are orthogonal wrt. the canonical scalar product
$X \lesssim Y$	there exists a possible parameter dependent constant $c > 0$ s.t. $X \leqslant cY$
$A \prec (\preccurlyeq) B$	the positive (semi-) definite partial order of matrices, i.e. $B - A \succ (\succcurlyeq) 0$ is positive (semi-) definite

(Differential-) Operators

$\frac{\partial}{\partial x}$	the partial derivative in the variable x
∇, ∇_x	the (row) vector of partial derivatives $\nabla = (\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_2})$
Tr	the trace operator of a matrix A, $\mathrm{Tr}\, A = \sum_j A_{jj}$ (sum of diagonal entries)
$\mathbf{Tr}$	vectorized version of Tr (a tensor contraction)
div_x	the divergence operator of a vector field v, $\mathrm{div}_x\, v = \mathrm{Tr}(\nabla_x \otimes v) = \sum_j \frac{\partial}{\partial x_j} v_j$
$\mathbf{div}_x$	vectorized version of div_x applied column-wise to a matrix $\mathrm{div}\, A = (\mathrm{div}\, A_{\cdot 1}, \mathrm{div}\, A_{\cdot 2}, \cdots)$
$\ln^+$	the positive part of the logarithm, $\ln^+ x = \max\{\ln(x), 0\}$.
$\mathcal{P}, \mathcal{L}$	Markov semigroup and its infinitesimal generator
A^*	the adjoint or dual operator to A, which is the transpose in the case of matrices or vectors.
$\phi_\# \mu$	push forward of a measure μ under the mapping ϕ, i.e. $\phi_\# \mu(A) = \mu(\phi^{-1}(A))$.
$\mathscr{D}_\theta X$	the (Malliavin–/L^p–) directional derivative of X in the direction of θ.

Index